OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: August 18, 2003, 19:53:48 ; Search time 4471.78 Seconds
                                    (without alignments)
                                    9743.047 Million cell updates/sec

Title: US-09-987-455-7
Perfect score: 1065
Sequence: 1 tctgagggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : GenEmbl:*
      1: gb_ba:*
      2: gb_htg:*
      3: gb_in:*
      4: gb_om:*
      5: gb_ov:*
      6: gb_pat:*
      7: gb_ph:*
      8: gb_pl:*
      9: gb_pr:*
     10: gb_ro:*
     11: gb_sts:*
     12: gb_sy:*
     13: gb_un:*
     14: gb_vi:*
     15: em_ba:*
     16: em_fun:*
     17: em_hum:*
     18: em_in:*
     19: em_mu:*
     20: em_om:*
     21: em_or:*
     22: em_ov:*
     23: em_pat:*
     24: em_ph:*
     25: em_pl:*
     26: em_ro:*
     27: em_sts:*
     28: em_un:*
     29: em_vi:*
     30: em_htg_hum:*
     31: em_htg_inv:*
     32: em_htg_other:*
     33: em_htg_mus:*
     34: em_htg_pln:*
     35: em_htg_rod:*
     36: em_htg_mam:*
     37: em_htg_vrt:*
     38: em_sy:*
     39: em_htgo_hum:*
     40: em_htgo_mus:*
     41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1065	100.0	1065	6	AX431270	AX431270 Sequence
2	1065	100.0	1065	6	AX431273	AX431273 Sequence
3	1065	100.0	1068	6	A27431	A27431 DNA sequenc
4	1065	100.0	1068	6	AR059987	AR059987 Sequence
5	1065	100.0	1068	6	E01935	E01935 Synthetic D
6	1065	100.0	1170	6	A27727	A27727 DNA sequenc
7	1065	100.0	1170	6	E01945	E01945 Synthetic D
8	1065	100.0	1314	6	A27435	A27435 DNA sequenc
9	1065	100.0	1314	6	E01937	E01937 Synthetic D
10	1065	100.0	1419	6	A27725	A27725 DNA sequenc
11	1065	100.0	1419	6	AR059996	AR059996 Sequence
12	1065	100.0	1419	6	E01944	E01944 Synthetic D
13	1065	100.0	1689	6	I06609	I06609 Sequence 44
14	1065	100.0	1689	6	I06614	I06614 Sequence 46
15	1065	100.0	1689	6	I08789	I08789 Sequence 3
16	1065	100.0	1780	6	A06611	A06611 Synthetic n
17	1065	100.0	1780	6	A10226	A10226 Synthetic D
18	1065	100.0	1836	6	E01176	E01176 DNA encodin
19	1065	100.0	1836	6	I01581	I01581 Sequence 1
20	1065	100.0	1981	6	I01583	I01583 Sequence 3
21	1065	100.0	1989	6	I01256	I01256 Sequence 2
22	1065	100.0	2099	6	A07207	A07207 Artificial
23	1065	100.0	2100	6	A07197	A07197 H.sapiens m
24	1065	100.0	2100	6	A27429	A27429 cDNA sequen
25	1065	100.0	2100	6	E01934	E01934 DNA encodin
26	1065	100.0	2101	6	AR059986	AR059986 Sequence
27	1065	100.0	2151	6	E00654	E00654 cDNA encodi
28	1065	100.0	2162	6	AR017909	AR017909 Sequence
29	1065	100.0	2162	6	AR017931	AR017931 Sequence
30	1065	100.0	2162	6	I18796	I18796 Sequence 3
31	1065	100.0	2249	6	A07209	A07209 Artificial
32	1065	100.0	2544	6	I08530	I08530 Sequence 1
33	1065	100.0	2547	6	E01163	E01163 cDNA encodi
34	1065	100.0	2547	6	I07841	I07841 Sequence 1
35	1065	100.0	2547	6	I07989	I07989 Sequence 1
36	1065	100.0	2547	6	I09622	I09622 Sequence 1
37	1065	100.0	2549	6	E01055	E01055 DNA encodin
38	1065	100.0	2560	6	A01465	A01465 H.sapiens t
39	1065	100.0	7533	6	I08638	I08638 Sequence 2
40	1064	99.9	1689	6	I08637	I08637 Sequence 1
41	1063.4	99.8	1170	6	AR059997	AR059997 Sequence
42	1063.4	99.8	1314	6	AR059989	AR059989 Sequence
43	1063.4	99.8	1506	6	E02247	E02247 cDNA sequen
44	1063.4	99.8	1680	6	I06422	I06422 Sequence 40
45	1063.4	99.8	1689	6	A14927	A14927 tPA-DNA fro

ALIGNMENTS

RESULT 1
AX431270
LOCUS   AX431270   1065 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 4 from Patent WO0240650.
ACCESSION   AX431270
VERSION   AX431270.1 GI:21656152
KEYWORDS
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
            1
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 4 23-MAY-2002;

Pred. No. is the number of results predicted by chance to have a

```
Qy    901  TGCCAGGGCGATTCGGAGGCCCCCTGTGTGTCTGAACGATGCCGATGACTTTGGTG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TGCCAGGGCGATTCGGAGGCCCCCTGTGTGTCTGAACGATGCCGATGACTTTGGTG  960

Qy    961  GGGCATCATCAGCTGGGGCGCTGGGGCTGGGCGTGCAGCAGAAGGATGTCCCGGGTGTGTACCACAAAG  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GGGCATCATCAGCTGGGGCGCTGGGGCTGGGCGTGCAGCAGAAGGATGTCCCGGGTGTGTACCACAAAG  1020

Qy   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA  1065
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA  1065
```

RESULT 2
AX431273

LOCUS   AX431273   1065 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION   Sequence 7 from Patent WO0240650.
ACCESSION   AX431273
VERSION   AX431273.1 GI:21655155
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
             or k2s molecules
  JOURNAL   Patent: WO 0240650-A 7 23-MAY-2002;
             BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
     source   1..1065
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for K2S protein"
BASE COUNT   225 a   314 c   312 g   214 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1065;
Best Local Similarity 100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  60

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120

Qy    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  180

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG  240

Qy    241  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGCTACAGCAGCCAGCCTCAG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGCTACAGCAGCCAGCCTCAG  300

Qy    301  TTTCGCATCAAAGAGGGCTCTTCGCCGACATCGCCTCCACCCTGCCAGGCTGCCATC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TTTCGCATCAAAGAGGGCTCTTCGCCGACATCGCCTCCACCCTGCCAGGCTGCCATC  360

Qy    361  TTTGCCAAGCACAGGAGGTCGCCCGCACTGGGCGTTCCGTGCGGGCATATTACTGCCGG  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TTTGCCAAGCACAGGAGGTCGCCCGCACTGGGCGTTCCGTGCGGGCATATTACTGCCGG  420

Qy    421  TCCTGCTGGATTCCTCTGCCCCACTGG  ...
```

```
Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  60

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120

Qy    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  180

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG  240
```

BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
     source   1..1065
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for K2S protein"
BASE COUNT   225 a   314 c   312 g   214 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1065;
Best Local Similarity 100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Wed Aug 20 09:29:43 2003   us-09-987-455-7.rge   Page 3

```
Db      4  TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC   63

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCTCCCGTGCTCCCGTGCTCCCGTGGGTGCCTCCCGTGCTCCCGTGGGTGCCTCGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64  CTCACCGAGTCGGGTGCCTCCTGCTCCCGTGCTCCCGTGGGTGCCTCGATCCTGATAGGCAAGGTT  123

Qy    121  TACACAGCACAGAACCCAGTGCCCAGGCACTGGCCTTGGGCAAACATAATTACTGCCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124  TACACAGCACAGAACCCAGTGCCCAGGCACTGGCCTGGGCAAACATAATTACTGCCGG  183

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCTGAAGAACCGAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||
Db    184  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG  243

Qy    241  GAGTACTGTGATGTGCCCTGCTCCACCTGCGCCTGACGAGTACAGCCAGCCTCAG  300
           ||||||||||||||||||||||||||||||||||||||||
Db    244  GAGTACTGTGATGTGCCCTGCTCCACCTGCGCCTGAGACACTACAGCCAGCCTCAG  303

Qy    301  TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCGGCCTCCCACCCTGGCAGGCTGCCATC  360
           ||||||||||||||||||||||||||||||||||||||||
Db    304  TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCGGCCTCCCACCCTGGCAGGCTGCCATC  363

Qy    361  TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGGTTCCTGTGCGGGGGCATACTCATCAGC  420
           ||||||||||||||||||||||||||||||||||||||||
Db    364  TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGGTTCCTGTGCGGGGGCATACTCATCAGC  423

Qy    421  TCCTGCTGATTCTCTCTGCCGACTGCTTCGAGGAGGTTCCGCCCACCACCTG  480
           ||||||||||||||||||||||||||||||||||||||||
Db    424  TCCTGCTGATTCTCTCTGCCGACTGCTTCGAGGAGGTTCCGCCCACCACCTG  483

Qy    481  ACGGTGATCTTGGCAGAACATACCGGTGGTCCTGGCGTGGTCCTGGCGAGGAGCAGAAATTTGAA  540
           ||||||||||||||||||||||||||||||||||||||||
Db    484  ACGGTGATCTTGGCAGAACATACCGGTGGTCCTGGCGTGGTCCTGGCGAGGAGCAGAAATTTGAA  543

Qy    541  GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTCGACAATGACATTGCG  600
           ||||||||||||||||||||||||||||||||||||||||
Db    544  GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTCGACAATGACATTGCG  603

Qy    601  CTGCTGCAGCTGAAATCGGATCGCCCGCCGGACTGAGCTGCGGACTGAGCTGGTGGTCCGCACT  660
           ||||||||||||||||||||||||||||||||||||||||
Db    604  CTGCTGCAGCTGAAATCGGATCGCCCGCCGGACTGAGCTGCGGACTGAGCTGGTGGTCCGCACT  663

Qy    661  GTGTGCCTTCCCCGCCGGCGGACCTGCAGCTGCAGCTGGAGCTGAAGGAGAGCTCATGTC  720
           ||||||||||||||||||||||||||||||||||||||||
Db    664  GTGTGCCTTCCCCGCCGGCGGACCTGCAGCTGCAGCTGGAGCTGAAGGAGAGCTCATGTC  723

Qy    721  TACGGCAAGCATGAGGCCTTGTCTCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  780
           ||||||||||||||||||||||||||||||||||||||||
Db    724  TACGGCAAGCATGAGGCCTTGTCTCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  783

Qy    781  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACGAC  840
           ||||||||||||||||||||||||||||||||||||||||
Db    784  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACGAC  843

Qy    841  AACATGCTGTGTGCTGGACACTCGGAGCGGCGGCGGCCCCAGGCAAACTTGCACGAGCGCC  900
           ||||||||||||||||||||||||||||||||||||||||
Db    844  AACATGCTGTGTGCTGGACACTCGGAGCGGCGGCGGCCCCAGGCAAACTTGCACGAGCGCC  903

Qy    901  TGCCAGGGCGATTCGGGAGGCCCCGTGGTGTCTGAACGATGCCGCATGAGACTTTGGTG  960
           ||||||||||||||||||||||||||||||||||||||||
Db    904  TGCCAGGGCGATTCGGGAGGCCCCGTGGTGTCTGAACGATGCCGCATGAGACTTTGGTG  963

Qy    961  GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1020
           ||||||||||||||||||||||||||||||||||||||||
Db    964  GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1023

Qy   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1065

Db   1024  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1068
```

RESULT 3
A27431
LOCUS   A27431   1068 bp   DNA   linear   PAT 09-JUL-2002
DEFINITION DNA sequence of coding regionin pTTkPAdeltatrp.
ACCESSION   A27431
VERSION   A27431.1   GI:2172228
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 31 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES   Location/Qualifiers
     source   1..1068
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   226 a   314 c   313 g   215 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1068;
Best Local Similarity   100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1  TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC   60
           ||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
AR059987

LOCUS   AR059987   1068 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 44 from patent US 5840533.
ACCESSION   AR059987
VERSION   AR059987.1 GI:5986437
KEYWORDS   Unknown.
SOURCE   Unknown.
  ORGANISM Unclassified.
REFERENCE   1 (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 44 24-NOV-1998;
FEATURES   Location/Qualifiers
     source   1..1068
                     /organism="unknown"
BASE COUNT   226 a   314 c   313 g   215 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1068;
Best Local Similarity   100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTGGCACGCACAGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTGGCACGCACAGC 63

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCTCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 CTCACCGAGTCGGGTGCCTCCTGCTCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT 123

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 183

Qy   181 AATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCGCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 AATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCGCACGTGCTGAAGAACCGCAGGCTGACGTGG 243

Qy   241 GAGTACTGTGATGTGCCCCTCCTGCGGCCTCCACCTGCGGCCTCCGAGACAGTACAGCCAGCCTCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 GAGTACTGTGATGTGCCCTCCTGCGGCCTCCACCTGCGGCCTCCGAGACAGTACAGCCAGCCTCAG 303

Qy   301 TTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 TTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 363

Qy   361 TTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCATACTCATCAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCATACTCATCAGC 423

Qy   421 TCCTGCTCGGATTCTCTCGCCGCCACTGCTTCCGCGCCACTGCTTCCGCCCCCCACCACCTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 TCCTGCTCGGATTCTCTCGCCGCCACTGCTTCCGCGCCACTGCTTCCGCCCCCCACCACCTG 483

Qy   481 ACGGTGATCTTGGGCAGAACATACGGGGTGCCTGGCGAGGAGCGACAATGACATTGCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 ACGGTGATCTTGGGCAGAACATACGGGGTGCCTGGCGAGGAGCGACAATGACATTGCG 543

Qy   541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACCAATGACATTGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACCAATGACATTGCG 603

Qy   601 CTGCTCGCAGCTGAAATCGGATTCGTCCCCGTGTCCCAGGAGAGCAGCGCCTGGTTCCGCACT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604 CTGCTCGCAGCTGAAATCGGATTCGTCCCCGTGTCCCAGGAGAGCAGCGCCTGGTTCCGCACT 663

Qy   661 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACTGACCGGAGTGTGAGCTCTCCGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACTGACCGGAGTGTGAGCTCTCCGGC 723

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   724 TACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 783

Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 843

Qy   841 AACATGCTGTGTCTGAGACACTCGGAGCGGGGGCGGCGGGGGCAAACTTGCACGACGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844 AACATGCTGTGTCTGAGACACTCGGAGCGGGGGCGGCGGGGGCAAACTTGCACGACGCC 903

Qy   901 TGCCAGGGCGATTCGGGAGGCCCCTGGTGTGTCTGACGATGGCCGATGACTTTGGTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   904 TGCCAGGGCGATTCGGGAGGCCCCTGGTGTGTCTGACGATGGCCGATGACTTTGGTG 963

Qy   961 GGCATCATCAGCTGGGCGCTGGGCCTGGGCTGTGACGACGAAGGATGTCCCGGGTGTGTACACAAAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964 GGCATCATCAGCTGGGCGCTGGGCCTGGGCTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1023

Qy   1021 GTTACCAACTACCTAGACTGAATTCGTGACAACATGCGACCGTGA 1065
              ||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GTTACCAACTACCTAGACTGAATTCGTGACAACATGCGACCGTGA 1068

RESULT 5
E01935
LOCUS   E01935   1068 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding new tissue plasminogen activator (t-PA).
ACCESSION   E01935
VERSION   E01935.1 GI:2170183
KEYWORDS   JP 1989104167-A/2.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 2 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT   OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/2
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19)
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   *source: clone=pTTkPAdeltatrp;
            FH   Key   Location/Qualifiers
            FH
            FT   CDS   1..1068
            FT   /product='t-PA'.
FEATURES   Location/Qualifiers
     source   1..1068
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   226 a   314 c   313 g   215 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1068;
Best Local Similarity   100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTACCGTGCGCACCACAGC 60

```
A27727
LOCUS       A27727          1170 bp    DNA        linear    PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pmTTk.
ACCESSION   A27727
VERSION     A27727.1  GI:21727237
KEYWORDS    
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1170)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 51 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1170
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      249 a     338 c     346 g     237 t
ORIGIN

Query Match            100.0%;  Score 1065;  DB 6;  Length 1170;
Best Local Similarity  100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC    165

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATCCTGATAGGCAAGGTT          120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATCCTGATAGGCAAGGTT         225

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCGACTGGCCTGGGCAAACATAATTACTGCCGG    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 TACACAGCACAGAACCCCAGTGCCAGGCGACTGGCCTGGGCAAACATAATTACTGCCGG    285

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG    345

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCCTGGGCCTGAGACAGTACAGCCTCAG    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCCTGGGCCTGAGACAGTACAGCCTCAG    405

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCCTCCTCACCCTGGCAGGCTGCCATC    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCCTCCTCACCCTGGCAGGCTGCCATC    465

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGCGGTTCCTGTGCGGGGCATACTCATCAGC      420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 TTTGCCAAGCACAGGAGGTCGCCCGGAGCGGTTCCTGTGCGGGGCATACTCATCAGC     525

Qy    421 TCCTGCTGTGGATTCTCTCTGCCGCGCCACTGCTTCCGGAGGAGGTTTCCGCCCCACCTG    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 TCCTGCTGTGGATTCTCTCTGCCGCGCCACTGCTTCCGGAGGAGGTTTCCGCCCCACCTG    585

Qy    481 ACGGTGATCTTGGGGCAGAGAACATACCGGGGTGGTCCCTGCGGAGGAGCAGCAGAAGTT    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 ACGGTGATCTTGGGGCAGAGAACATACCGGGGTGGTCCCTGCGGAGGAGCAGCAGAAGTT    645

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACTTACGACAATGACATTGCG    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACTTACGACAATGACATTGCG    705

Qy    601 CTGCTCGAGCTGAAATGGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTTCCGCCACT    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 CTGCTCGAGCTGAAATGGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTTCCGCCACT    765

Qy    661 GTGTGCCTTCCCCGGGCGACTCGCAGCTGCGACCGGAGTGTGAGCTCTCCGGC         720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      4 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC    63
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATCCTGATAGGCAAGGTT         120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATCCTGATAGGCAAGGTT        123

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCGACTGGCCTGGGCAAACATAATTACTGCCGG    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 TACACAGCACAGAACCCCAGTGCCAGGCGACTGGCCTGGGCAAACATAATTACTGCCGG    183

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG    243

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCCTGGGCCTGAGACAGTACAGCCTCAG    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCCTGGGCCTGAGACAGTACAGCCTCAG    303

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCCTCCTCACCCTGGCAGGCTGCCATC    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCCTCCTCACCCTGGCAGGCTGCCATC    363

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGCGGTTCCTGTGCGGGGCATACTCATCAGC      420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TTTGCCAAGCACAGGAGGTCGCCCGGAGCGGTTCCTGTGCGGGGCATACTCATCAGC     423

Qy    421 TCCTGCTGTGGATTCTCTCTGCCGCGCCACTGCTTCCGGAGGAGGTTTCCGCCCCACCTG    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 TCCTGCTGTGGATTCTCTCTGCCGCGCCACTGCTTCCGGAGGAGGTTTCCGCCCCACCTG    483

Qy    481 ACGGTGATCTTGGGGCAGAGAACATACCGGGGTGGTCCCTGCGGAGGAGCAGCAGAAGTT    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 ACGGTGATCTTGGGGCAGAGAACATACCGGGGTGGTCCCTGCGGAGGAGCAGCAGAAGTT    543

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACTTACGACAATGACATTGCG    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACTTACGACAATGACATTGCG    603

Qy    601 CTGCTCGAGCTGAAATGGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTTCCGCCACT    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 CTGCTCGAGCTGAAATGGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTTCCGCCACT    663

Qy    661 GTGTGCCTTCCCCGGGCGACTCGCAGCTGCGACCGGAGTGTGAGCTCTCCGGC         720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 GTGTGCCTTCCCCGGGCGACTCGCAGCTGCGACCGGAGTGTGAGCTCTCCGGC        723

Qy    721 TACGGCAAGCAATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724 TACGGCAAGCAATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC    783

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACGAACAGTCACCGAC    840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACGAACAGTCACCGAC    843

Qy    841 AACATGCTGTGTGCTGGAGACACTCGAGGCGGCGGCCCCCAGGCAAACTTGCAGACGCGCC    900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 AACATGCTGTGTGCTGGAGACACTCGAGGCGGCGGCCCCCAGGCAAACTTGCAGACGCGCC    903

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG    963

Qy    961 GGCATCATCAGCTGGGGGCCTGTGGACAGAAGATGTCCCGGGTGTGTACACAAAG      1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 GGCATCATCAGCTGGGGGCCTGTGGACAGAAGATGTCCCGGGTGTGTACACAAAG     1023

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGAACCGTGA     1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGAACCGTGA     1068
```

RESULT 6

```
Query Match            100.0%;  Score 1065;  DB 6;  Length 1170;
Best Local Similarity  100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGCACGCACAGC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGCACGCACAGC  165

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  225

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTCGGCCAAACATAATTACTGCCGG  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTCGGCCAAACATAATTACTGCCGG  285

Qy    181 AATCCTGATGGGGATGCCAAGCCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 AATCCTGATGGGGATGCCAAGCCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  345

Qy    241 GAGTACTGTGATGTGCCCTCTGCTCACCTGGGCTGAGCAGTACAGCCAGCCTCAG  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 GAGTACTGTGATGTGCCCTCTGCTCACCTGGGCTGAGCAGTACAGCCAGCCTCAG  405

Qy    301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCGAGGCTGCCATC  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCGAGGCTGCCATC  465

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGGAGGTTCCTGTGCGCCGGAGAGGTTCCGCGCCCACCTG  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 TTTGCCAAGCACAGGAGGTCGCCCGGAGGAGGTTCCTGTGCGCCGGAGAGGTTCCGCGCCCACCTG  525

Qy    421 TCCTGCTGGATTCTCTCTGCGCCCACTGTCTTCCCAGGAGGAGGTTTCCGCCCCCACCACCTG  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 TCCTGCTGGATTCTCTCTGCGCCCACTGTCTTCCCAGGAGGAGGTTTCCGCCCCCACCACCTG  585

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGGAGCAGAAATTTGAA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGGAGCAGAAATTTGAA  645

Qy    541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  705

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTCCGCCACT  660
            ||||||||||||||||||||||||||||||||||||||||
Db    706 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTCCGCCACT  765

Qy    661 GTGTGCCTTCCCCGGCCGACTGAGCTCTCCGGC  720
            ||||||||||||||||||||||||||||||||||
Db    766 GTGTGCCTTCCCCGGCCGACTGAGCTCTCCGGC  825

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTGAGGAGGCTCATGTC  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTGAGGAGGCTCATGTC  885

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  945

Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCCGGCGGCGGCCCCAGGCAAACTTGCACGACGCC  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 AACATGCTGTGTCTGGAGACACTCGGAGCGGCCGGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1005

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCAACGATGCCGCATGACTTTGGTG  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCAACGATGCCGCATGACTTTGGTG 1065

Qy    961 GGCATCATCAGCTGGGCCTGGGGCGGGGTGTGTGACAGAAGGATGTCCCGGGTGTACACAAAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 GGCATCATCAGCTGGGCCTGGGGCGGGGTGTGTGACAGAAGGATGTCCCGGGTGTACACAAAG 1125

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
            |||||||||||||||||||||||||||||||||||||||||||||
Db   1126 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1170
```

```
Db    766 GTGTGCCTTCCCCGGCCGACCTGCAGCTGCCGGAGCTGCCGGAGCGGAGTGTGAGCGGTCTCCGGC  825
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTGAGGAGGCTCATGTC  780

Db    826 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTGAGGAGGCTCATGTC  885
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840

Db    886 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  945
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCCGGCGGCGGCCCCAGGCAAACTTGCACGACGCC  900

Db    946 AACATGCTGTGTCTGGAGACACTCGGAGCGGCCGGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1005
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGATGGCCGCATGACTTTGGTG  960

Db   1006 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGATGGCCGCATGACTTTGGTG 1065
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961 GGCATCATCAGCTGGGCCTGGGGCGGGGTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020

Db   1066 GGCATCATCAGCTGGGCCTGGGGCGGGGTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1125
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065

Db   1126 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1170
```

```
RESULT 7
E01945
LOCUS       E01945                   1170 bp    DNA    linear       PAT 29-SEP-1997
DEFINITION  Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01945
VERSION     E01945.1  GI:2170193
KEYWORDS    JP 1989104167-A/12.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1170)
  AUTHORS   Niwa.M., Saito,Y., Sasaki.H., Hayashi,M., Notani,J. and
            Kobayashi.M.
  TITLE     NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 12 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/12
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19),
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pmTTk;
            FH   Key           Location/Qualifiers
            FH
            FT   CDS              1..1170
            FT                    /product='new t-PA'.
FEATURES             Location/Qualifiers
     source          1..1170
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT  249 a   338 c   346 g   237 t
ORIGIN
```

RESULT 8
A27435
LOCUS   A27435   1314 bp   DNA   linear   PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pTQkPAdeltatrp.
ACCESSION   A27435
VERSION   A27435.1 GI:21727230
KEYWORDS
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 35 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES   Location/Qualifiers
     source   1..1314
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   286 a   386 c   393 g   249 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1314;
Best Local Similarity   100.0%;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGC 309

Qy   61 CTCACCGAGTCGGGTGCCCTGCCTCCCGTGCAATTCCATGATCCTGATAGGCAAGGTT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   310 CTCACCGAGTCGGGTGCCCTGCCTCCCGTGCAATTCCATGATCCTGATAGGCAAGGTT 369

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCGCTGGGCAAACATAATTACTGCCGG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   370 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCGCTGGGCAAACATAATTACTGCCGG 429

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGAGCCTGACGTGG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGAGCCTGACGTGG 489

Qy   241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGCTGAGCAGCCAGCCTCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   490 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGCTGAGCAGCCAGCCTCAG 549

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCGGGTTCCTCCACCCCTGGCAGGCTGCCATC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   550 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCGGGTTCCTCCACCCCTGGCAGGCTGCCATC 609

Qy   361 TTTGCCAAGCACACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   610 TTTGCCAAGCACACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 669

Qy   421 TCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCCGCGGAGAGGTTTCCGCCCCACCACCTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   670 TCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCCGCGGAGAGGTTTCCGCCCCACCACCTG 729

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCGAGGAGCAGCAGAAATTTGAA 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   730 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCGAGGAGCAGCAGAAATTTGAA 789

Qy   541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACAATGACATTGCG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   790 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACAATGACATTGCG 849

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGGAGCAGCGTGGTCCGCACT 660

(alignment continues)

Db   1126 GTTACCAACTACCTAGACTGGATTCGTGTCGTGACAACATGCCAACCGTGA 1170

RESULT 9
E01937
LOCUS   E01937   1314 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01937
VERSION   E01937.1 GI:2170185
KEYWORDS   JP 1989104167-A/4.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences
REFERENCE   1 (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 4 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT   OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/4
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
                 C12R1:19),
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pTQkPAdeltatrp;
FEATURES   Location/Qualifiers
FH   Key   Location/Qualifiers
FT   CDS   1..1314
FT   /product='t-PA'.
     source   1..1314
                     /organism="synthetic construct"

```
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT      286 a    386 c    393 g    249 t
ORIGIN

Query Match             100.0%;  Score 1065;  DB 6;  Length 1314;
Best Local Similarity   100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  250  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC   309

Qy   61  CTCACCGAGTCGGGTGCCCTCCGTGCCTCCCGTGATCCTGATAGGCAAGGTT        120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  310  CTCACCGAGTCGGGTGCCCTCCGTGCCTCCCGTGATCCTGATAGGCAAGGTT        369

Qy  121  TACACAGCACAGAACCCCAGTGCCGACTGGGCCTGGCAAACATAATTACTGCGG       180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  370  TACACAGCACAGAACCCCAGTGCCGACTGGGCCTGGCAAACATAATTACTGCGG       429

Qy  181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  430  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  489

Qy  241  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGCCGACATGCATGCGCCAGCCAGCCTCAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  490  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGCCGACATGCATGCGCCAGCCAGCCTCAG 549

Qy  301  TTTCGCATCAAAGGAGGGCTCTTCGCGGACATCGCCGACATCGCCGGCTGCCATC       360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  550  TTTCGCATCAAAGGAGGGCTCTTCGCGGACATCGCCGACATCGCCGGCTGCCATC       609

Qy  361  TTTGCCAAGCACAGGAGGTCGCCGGAGAGCGTTCGCCGGCCCACTGCTTCCGAGGAGAGG  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  610  TTTGCCAAGCACAGGAGGTCGCCGGAGAGCGTTCGCCGGCCCACTGCTTCCGAGGAGAGG  669

Qy  421  TCCTGCTGGATTCTCTGCTGCCGCCGGCCACTCCGGGTGGTGAGCGGCGAGAGAGG      480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  670  TCCTGCTGGATTCTCTGCTGCCGCCGGCCACTCCGGGTGGTGAGCGGCGAGAGAGG      729

Qy  481  ACGGTGATCTTGGGCAGAGAACATACCGGGTGGTCCTGCTGCCAGAAATTTGAA        540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  730  ACGGTGATCTTGGGCAGAGAACATACCGGGTGGTCCTGCTGCCAGAAATTTGAA        789

Qy  541  GTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  790  GTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   849

Qy  601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTTCCGCACT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  850  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTTCCGCACT 909

Qy  661  GTGTGCCTTCCCCCGGCGCAGCTGCCGACTGCTGAGCGTGTGAGCTGTTCGAGCTGTTCCGGC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  910  GTGTGCCTTCCCCCGGCGCAGCTGCCGACTGCTGAGCGTGTGAGCTGTTCGAGCTGTTCCGGC 969

Qy  721  TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCGGCCTGAAGGAGGCTCATGTC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  970  TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCGGCCTGAAGGAGGCTCATGTC 1029

Qy  781  AGACTGTACCCATCCAGCCGCTGCACATCAACAACAACATTTACTTAACGACAACAGTCACCGAC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1030  AGACTGTACCCATCCAGCCGCTGCACATCAACAACAACATTTACTTAACGACAACAGTCACCGAC 1089

Qy  841  AACATGCTGTGTGCTGGACACTCGGAGGCGGGGCCCCCCAGGCCAAACTTGACGACGCC     900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1090  AACATGCTGTGTGCTGGACACTCGGAGGCGGGGCCCCCCAGGCCAAACTTGACGACGCC     1149

Qy  901  TGCCAGGGCGATTCGGGAGGGCCCCCCGTGTGTGTCTGAACGATGCCGCATGAGACTTTGGTG   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1150  TGCCAGGGCGATTCGGGAGGGCCCCCCGTGTGTGTCTGAACGATGCCGCATGAGACTTTGGTG   1209
```

```
Qy  961  GGCATCATCAGCTGGGCCTGGGCCTGGGCCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1210  GGCATCATCAGCTGGGCCTGGGCCTGGGCCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG  1269

Qy 1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA                     1065
         |||||||||||||||||||||||||||||||||||||||||||||
Db 1270  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA                     1314
```

```
RESULT 10
LOCUS       A27725                  1419 bp    DNA     linear   PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pmTQk112.
ACCESSION   A27725
VERSION     A27725.1  GI:2172236
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 49 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1419
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      309 a    413 c    426 g    271 t
ORIGIN

Query Match             100.0%;  Score 1065;  DB 6;  Length 1419;
Best Local Similarity   100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  355  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC   414

Qy   61  CTCACCGAGTCGGGTGCCCTCCGTGCCTCCCGTGATCCTGATAGGCAAGGTT        120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  415  CTCACCGAGTCGGGTGCCCTCCGTGCCTCCCGTGATCCTGATAGGCAAGGTT        474

Qy  121  TACACAGCACAGAACCCCAGTGCCGACTGGGCCTGGCAAACATAATTACTGCGG       180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  475  TACACAGCACAGAACCCCAGTGCCGACTGGGCCTGGCAAACATAATTACTGCGG       534

Qy  181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  535  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  594

Qy  241  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGCCGACATGCATGCGCCAGCCAGCCTCAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  595  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGCCGACATGCATGCGCCAGCCAGCCTCAG 654

Qy  301  TTTCGCATCAAAGGAGGGCTCTTCGCGGACATCGCCGACATCGCCGGCTGCCATC       360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  655  TTTCGCATCAAAGGAGGGCTCTTCGCGGACATCGCCGACATCGCCGGCTGCCATC       714

Qy  361  TTTGCCAAGCACAGGAGGTCGCCGGAGAGCGTTCGCCGGCCCACTGCTTCCGAGGAGAGC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  715  TTTGCCAAGCACAGGAGGTCGCCGGAGAGCGTTCGCCGGCCCACTGCTTCCGAGGAGAGC  774

Qy  421  TCCTGCTGATTCTCTGCTGCCGCCCAGCCCCAGGCCAAACTTGACGACGCC          480
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  775  TCCTGCTGATTCTCTGCTGCCGCCCAGCCCCAGGCCAAACTTGACGACGCC          834

Qy  481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCTGCCGGAGAGCAGAGAAATTTGAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  835  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCTGCCGGAGAGCAGAGAAATTTGAA 894
```

Qy   541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 954

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 1014

Qy   661 GTGTGCCTTCCCCGGCGACCTGCAGCTGCGACCTGGAGCGGAGTGTGAGCTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 GTGTGCCTTCCCCGGCGACCTGCAGCTGCGACCTGGAGCGGAGTGTGAGCTCTCCGGC 1074

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGAACGGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGAACGGCGGCTGAAGGAGGCTCATGTC 1134

Qy   781 AGACTGTACCCATCCAGCCGCTGGACACTCGGAGACACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 AGACTGTACCCATCCAGCCGCTGGACACTCGGAGACACAGTCACCGAC 1194

Qy   841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCCGGGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCCGGGGCAAACTTGCACGACGCC 1254

Qy   901 TGCCAGGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCCGCATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 TGCCAGGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCCGCATGACTTTGGTG 1314

Qy   961 GGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAGAAGGATGTCCCGGGTGTGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 GGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAGAAGGATGTCCCGGGTGTGTGTGTACACAAAG 1374

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||
Db   1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1419

RESULT 11
AR059996
LOCUS   AR059996   1419 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION   Sequence 62 from patent US 5840533.
ACCESSION   AR059996
VERSION   AR059996.1   GI:5986446
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
             Kobayashi,M.
  TITLE   Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 62 24-NOV-1998;
FEATURES   Location/Qualifiers
     source   1..1419
                     /organism="unknown"
BASE COUNT   309 a   416 c   423 g   271 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1419;
Best Local Similarity 100.0%;
Matches 1065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC. 414

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCCGTTGCCTCCGTCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCCGTTGCCTCCGTCCGTGAATTCCATGATCCTGATAGGCAAGGTT 474

Qy   121 TACAGCACAGACAGACCCCAGTGCCCAGGCACTGGGCCTGGCACAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 TACAGCACAGACAGACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 534

Qy   181 AATCCTGATGGGATGCCAAGCCCTGGTGCTGCTGAAGAACCGCAGGCTGCTGACGTGG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 AATCCTGATGGGATGCCAAGCCCTGGTGCTGCTGAAGAACCGCAGGCTGCTGACGTGG 594

Qy   241 GAGTACTGTGATGTGCCCTCCTGCTCCACTGCGCGCCTGAGACAGTACAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 GAGTACTGTGATGTGCCCTCCTGCTCCACTGCGCGCCTGAGACAGTACAGCCAGCCTCAG 654

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACCATGCGCCTCCACCCCTGGCGGGGCTGCCATC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   655 TTTCGCATCAAAGGAGGGCTCTTCGCCGACCATGCGCCTCCACCCCTGGCGGGGCTGCCATC 714

Qy   361 TTTGCCAAGCACAGGAGTCGCCCGGAGCGGTTCCTGTGCGGGGGCCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715 TTTGCCAAGCACAGGAGTCGCCCGGAGCGGTTCCTGTGCGGGGGCCATACTCATCAGC 774

Qy   421 TCCTGCTGATTCTCTGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   775 TCCTGCTGATTCTCTGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 834

Qy   481 ACGGTGATCTTGGCAGAACATACCGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835 ACGGTGATCTTGGCAGAACATACCGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 894

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGACGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895 GTCGAAAAATACATTGTCCATAAGGAATTCGATGACGATGACACTTACGACAATGACATTGCG 954

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 1014

Qy   661 GTGTGCCTTCCCCGGCGACCTGCAGCTGCGACCTGGAGCGGAGTGTGAGCTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 GTGTGCCTTCCCCGGCGACCTGCAGCTGCGACCTGGAGCGGAGTGTGAGCTCTCCGGC 1074

Qy   721 TACGGCAAGCATGAGGCCTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 TACGGCAAGCATGAGGCCTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 1134

Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1194

Qy   841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCCCAGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCCCAGGCAAACTTGCACGACGCC 1254

Qy   901 TGCCAGGGCGGATTCGGGAGGCGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 TGCCAGGGCGGATTCGGGAGGCGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 1314

Qy   961 GGCATCATCAGCTGGGGCCTGGGCGGCTGGGGCCTGGGGCGTGTGGACAGAAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 GGCATCATCAGCTGGGGCCTGGGCGGCTGGGGCCTGGGGCGTGTGGACAGAAGAAGGATGTCCCGGGTGTGTACACAAAG 1374

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1419

RESULT 12
E01944
LOCUS   E01944   1419 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION   Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01944
VERSION   E01944.1   GI:2170192
KEYWORDS   JP 1989104167-A/11.
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1 (bases 1 to 1419)

```
AUTHORS      Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
             Kobayashi,M.
TITLE        NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
JOURNAL      Patent: JP 1989104167-A 11 21-APR-1989;
             FUJISAWA PHARMACEUT CO LTD
COMMENT      OS   Artificial gene
             OC   Artificial sequence; Genes.
             PN   JP 199104167-A/11
             PD   21-APR-1989
             PF   01-AUG-1988 JP 1988192320
             PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
             13-NOV-1987 GB 87 8726683
             PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
             PI   NOTANI JOJI,
             PI   KOBAYASHI MASAKAZU
             PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
             C12R1:19),
             PC   (C12N9/64,C12R1:91);
             CC   strandedness: Double;
             CC   topology: Linear;
             CC   hypothetical: No;
             CC   anti-sense: No;
             CC   *source: clone=pmTQk112;
             FH   Key             Location/Qualifiers
             FH
             FT   CDS             1..1419
             FT                   /product='new t-PA'.
FEATURES     source          1..1419
                             /organism="synthetic construct"
                             /mol_type="genomic DNA"
                             /db_xref="taxon:32630"
BASE COUNT     309 a    413 c    426 g    271 t
ORIGIN

Query Match             100.0%;  Score 1065;  DB 6;  Length 1419;
Best Local Similarity   100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    355 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 414

Qy     61 CTCACCGAGTCGGGTGCCTCTCGTGCCTCCCGTGCCTCCCGTGCCTGAATTCCATGATCCTGATAGGCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    415 CTCACCGAGTCGGGTGCCTCTCGTGCCTCCCGTGCCTCCCGTGCCTGAATTCCATGATCCTGATAGGCAAGGTT 474

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGCCCCAGTGCCCAGCCCAGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    475 TACACAGCACAGAACCCCAGTGCCCAGCCCCAGTGCCCAGCCCAGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 534

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGCTGACGTGG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    535 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGCTGACGTGG 594

Qy    241 GAGTACTGTGATGTGCCCCTCCTGTCTCCACCTGGGCCTGGGCCTGAGACAGTAGCCAGGCCTCAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    595 GAGTACTGTGATGTGCCCCTCCTGTCTCCACCTGGGCCTGGGCCTGAGACAGTAGCCAGGCCTCAG 654

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCGACATCGCCGACATGCCCATC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCGACATCGCCGACATGCCCATC 714

Qy    361 TTTGCCAAGCACAGAGAGTCGCCCGAGAGCCGGTTCCTGTCGCGGGGCATACTCATCAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715 TTTGCCAAGCACAGAGAGTCGCCCGAGAGCCGGTTCCTGTCGCGGGGCATACTCATCAGC 774

Qy    421 TCCTGCTGGATTCTCTCTGCCCGCCACTGCTTCCGCGCCCACCACCTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 TCCTGCTGGATTCTCTCTGCCCGCCACTGCTTCCGCGCCCACCACCTG 834

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    835 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA 894
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 954

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCAGCGTGGTCCGCACT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCAGCGTGGTCCGCACT 1014

Qy    661 GTGTGCCTTTCCCCGGGGGACTGCAGCTGCCGGACTGACTGAGTGTGAGCTCTCCGGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 GTGTGCCTTTCCCCGGGGGACTGCAGCTGCCGGACTGACTGAGTGTGAGCTCTCCGGC 1074

Qy    721 TACGGCAAGCATGAGGGCCTTGTCTCCTTCTATTCGAGGCGGCTGAACGGAGGCTCATGTC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 TACGGCAAGCATGAGGGCCTTGTCTCCTTCTATTCGAGGCGGCTGAACGGAGGCTCATGTC 1134

Qy    781 AGACTGTACCCATCCAGCGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 AGACTGTACCCATCCAGCGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1194

Qy    841 AACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 AACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCC 1254

Qy    901 TGGCAGGCGGATTCGGGAGGCGCCCCTGGTGTCGTGAACGATGCCGATGACTTTGGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 TGGCAGGCGGATTCGGGAGGCGCCCCTGGTGTCGTGAACGATGCCGATGACTTTGGTG 1314

Qy    961 GGCATCATCAGCTGGGCGTGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 GGCATCATCAGCTGGGCGTGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1374

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1419
```

```
RESULT 13
I06609
LOCUS       I06609            Sequence 44 from Patent EP 0293934          1689 bp    DNA          linear    PAT 02-DEC-1994
DEFINITION  Sequence 44 from Patent EP 0293934.
ACCESSION   I06609
VERSION     I06609.1  GI:590639
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown
            Unclassified.
REFERENCE   1  (bases 1 to 1689)
  AUTHORS   Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
            Hashimoto,A. and Yuzuriha,T.
  TITLE     Mutant t-PA with kringle replacement
  JOURNAL   Patent: EP 0293934-A1 44 07-DEC-1988;
FEATURES             Location/Qualifiers
     source          1..1689
                     /organism="unknown"
BASE COUNT    377 a    483 c    505 g    324 t
ORIGIN

Query Match             100.0%;  Score 1065;  DB 6;  Length 1689;
Best Local Similarity   100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 684

Qy     61 CTCACCGAGTCGGGTGCCTCTCGTGCCTCCCGTGCCTCCCGTGCCTGAATTCCATGATCCTGATAGGCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 CTCACCGAGTCGGGTGCCTCTCGTGCCTCCCGTGCCTCCCGTGCCTGAATTCCATGATCCTGATAGGCAAGGTT 744

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGCCCCAGTGCCCAGCCCAGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

REFERENCE   1 (bases 1 to 1689)
AUTHORS   Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
             Hashimoto,A. and Yuzuriha,T.
TITLE   Mutant t-PA with kringle replacement
JOURNAL   Patent: EP 0293934-A1 46 07-DEC-1988;
FEATURES   Location/Qualifiers
  source   1..1689
             /organism="unknown"

BASE COUNT   377 a   484 c   504 g   324 t
ORIGIN

Unclassified.

Query Match   100.0%; Score 1065; DB 6; Length 1689;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGTCAGCCTACCGTGCACGGCACAGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGTCAGCCTACCGTGCACGGCACAGC 684

Qy   61 CTCACCGAGTCGGGTGGCTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685 CTCACCGAGTCGGGTGGCTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 744

Qy   121 TACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   745 TACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 804

Qy   181 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 864

Qy   241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGTCCACCTGCTGAAGAACCGCAGGCTGACGTGG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGTCCACCTGCTGAAGAACCGCAGGCTGACGTGG 924

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCCACACATGCCTCCACCCTGCAGCTGCCATC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 TTTCGCATCAAAGGAGGGCTCTTCGCCCACACATGCCTCCACCCTGCAGCTGCCATC 984

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1044

Qy   421 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1045 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 1104

Qy   481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGGCAGAGCAGAAATTTGAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1105 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGGCAGAGCAGAAATTTGAA 1164

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1165 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1224

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGTGCCCAGGAGCAGCCTGGTCCGCACT 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1225 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGTGCCCAGGAGCAGCCTGGTCCGCACT 1284

Qy   661 GTGTGCCTTCCCCCGGCCGGACTCGGACTGGAGCTGTGAGCTCTCCGGC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1285 GTGTGCCTTCCCCCGGCCGGACTCGGACTGGAGCTGTGAGCTCTCCGGC 1344

Qy   721 TACGGCAAGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1345 TACGGCAAGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC 1404

Qy   781 AGACTGTACCCATCCAGCCGCCTGCACATTCACATTTAACAGAACAGTCACCGAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1405 AGACTGTACCCATCCAGCCGCCTGCACATCACAACATTTAACAGAACAGTCACCGAC 1464

Qy   841 AACATGCTGTGTCGTGGAGACACTCGGAGACTCGGAGCGGCGGCAAACTTGCACGACGCC 900

Db   745 TACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 804

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 864

Qy   241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGGGCCTGAGCAGCAGCCAGCCTCAG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGGGCCTGAGCAGCAGCCAGCCTCAG 924

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCCACACATGCCTCCACCCTGCAGGCTGCCATC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 TTTCGCATCAAAGGAGGGCTCTTCGCCCACACATGCCTCCACCCTGCAGGCTGCCATC 984

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1044

Qy   421 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1045 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 1104

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGCAGAGCAGAAATTTGAA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1105 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGCAGAGCAGAAATTTGAA 1164

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1165 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1224

Qy   601 GGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGAAGATGTCCCGGGGTGTGTACAAAG 1020

Db 1585 GGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGAAGATGTCCCGGGGTGTGTACAAAG 1644

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065

Db 1645 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1689

RESULT 14
I06614
LOCUS   I06614   1689 bp   DNA   linear   PAT 02-DEC-1994
DEFINITION Sequence 46 from Patent EP 0293934.
ACCESSION   I06614
VERSION   I06614.1 GI:590640
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.

```
Db   1465  AACATGTGTGTGTCTGGAGACACTCGGAGCGCGGGGCCCCCAGGCAAACTTGCGACGACGCC 1524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 1584
Qy    961  GGCATCATCAGCTGGGGCGCTGGGCGTGGGCGTGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585  GGCATCATCAGCTGGGGCGCTGGGCGTGGGCGTGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1644
Qy   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1645  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1689


RESULT 15
I08789

LOCUS       I08789                  1689 bp    DNA        linear    PAT 02-DEC-1994
DEFINITION  Sequence 3 from Patent WO 8804690.
ACCESSION   I08789
VERSION     I08789.1  GI:588494
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1689)
  AUTHORS   Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.
  JOURNAL   Patent: WO 8804690-A 3 30-JUN-1988;
FEATURES             Location/Qualifiers
     source          1..1689
                     /organism="unknown"
BASE COUNT      379 a    480 c    505 g    325 t
ORIGIN

Query Match      100.0%;     Score 1065;    DB 6;    Length 1689;
Best Local Similarity   100.0%;     Pred. No. 4.8e-248;
Matches 1065;    Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTCACCGTGGCACGCACAGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625  TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTCACCGTGGCACGCACAGC 684

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685  CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 744

Qy    121  TACCAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745  TACCAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 804

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCAGGCTGACGTGG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCAGGCTGACGTGG 864

Qy    241  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCCTCAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCCTCAG 924

Qy    301  TTTCGCATCAAAGGAGGGCTCTTCGCCCACACATCGCCCACACCCCTGCCATC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    925  TTTCGCATCAAAGGAGGGCTCTTCGCCCACACATCGCCCACACCCCTGCCATC 984

Qy    361  TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985  TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 1044

Qy    421  TCCTGCTGGATTCTCTTCTGCCGCCGCACCTGCTTCCAGGAGAGGTTTCCGCCCCACCTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045  TCCTGCTGGATTCTCTTCTGCCGCCGCACCTGCTTCCAGGAGAGGTTTCCGCCCCACCTG 1104

Qy    481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 1164

Qy    541  GTCGGAAAATACATTGTCCATAAGGAATTCGATGATGATGACAATGACATTGCG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165  GTCGGAAAATACATTGTCCATAAGGAATTCGATGATGATGACAATGACATTGCG 1224

Qy    601  CTGCTGCAGCTGAAATCGGATTCCTCCCGTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 1284

Qy    661  GTGTGCCTTCCCCGCCGGCGACCTGCAGCTGACCGACTGGACGGAGTGTGAGCTCTTCCGGC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285  GTGTGCCTTCCCCGCCGGCGACCTGCAGCTGACCGACTGGACGGAGTGTGAGCTCTCCGGC 1344

Qy    721  TACGGCAAGCAATGAGGCCTTGCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345  TACGGCAAGCAATGAGGCCTTGCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1404

Qy    781  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1464

Qy    841  AACATGTGTGTGTCTGGAGACACTCGGAGCGCGGGGCCCCCAGGCAAACTTGCGACGACGCC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465  AACATGTGTGTGTCTGGAGACACTCGGAGCGCGGGGCCCCCAGGCAAACTTGCGACGACGCC 1524

Qy    901  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 1584

Qy    961  GGCATCATCAGCTGGGGCGCTGGGCGTGGGCGTGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585  GGCATCATCAGCTGGGGCGCTGGGCGTGGGCGTGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1644

Qy   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1645  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1689


Search completed: August 19, 2003, 04:46:36
Job time : 4472.78 secs
```

OM nucleic - nucleic search, using sw model

```
          GenCore version 5.1.6
    Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: August 18, 2003, 19:51:53 ; Search time 375.192 Seconds
                                          (without alignments)
                                          7662.485 Million cell updates/sec

Title: US-09-987-455-7
Perfect score: 1065
Sequence: 1 tctgagggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0
Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_19Jun03;*
```
 1:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
 2:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
 3:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
 4:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
 5:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
 6:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
 7:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
 8:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
 9:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1065	100.0	1065	24	AAD40616	Human kringle 2 se
2	1065	100.0	1068	10	AAN91120	Sequence of coding
3	1065	100.0	1170	10	AAN91135	Sequence of coding
4	1065	100.0	1314	10	AAN91123	Sequence of coding
5	1065	100.0	1419	10	AAN91133	Sequence of coding
6	1065	100.0	1680	17	AAT27588	Novel plasminogen
7	1065	100.0	1680	17	AAT27587	Novel plasminogen
8	1065	100.0	1689	9	AAN82178	Tissue plasminogen
9	1065	100.0	1780	11	AAO01358	Sequence encoding
10	1065	100.0	2100	10	AAN91119	Sequence of native
11	1065	100.0	2100	11	AAO55532	Plasmid pST112 enc
12	1065	100.0	2162	16	AAO87370	Human tPA clone.
13	1065	100.0	2162	19	AAV37294	Human tissue plasm
14	1065	100.0	2547	7	AAN60659	Sequence encoding
15	1065	100.0	2560	10	AAN90542	Plasmid pKG12 cont
16	1065	100.0	7533	8	AAN70491	Entire sequence of
17	1065	100.0	7533	9	AAN81970	Plasmid pEMp1-tPA.
18	1063.4	99.8	1614	11	AAO03581	Glycosylated plasm
19	1063.4	99.8	1614	12	AAQ11550	Sequence encoding
20	1063.4	99.8	1659	12	AAQ11551	Sequence encoding
21	1063.4	99.8	1689	9	AAN82179	Tissue plasminogen
22	1063.4	99.8	1689	11	AAQ81090	Mutated t-PA analo
23	1063.4	99.8	1689	11	AAQ04903	Part of tPA024 gen
24	1063.4	99.8	1689	11	AAQ04904	Part of tPA023 gen
25	1063.4	99.8	1689	11	AAQ05177	Sequence encoding
26	1063.4	99.8	1689	12	AAQ13379	T-PA67+ mutant wit
27	1063.4	99.8	1689	24	ABK52546	Human tissue plasm
28	1063.4	99.8	1740	17	AAT27585	Full-length tissue
29	1063.4	99.8	1955	21	AAZ31998	Human tissue plasm
30	1063.4	99.8	1989	7	AAN60700	Sequence encoding
31	1063.4	99.8	2091	12	AAQ12073	T-PA variant contg
32	1063.4	99.8	2099	11	AAQ05534	Plasmid pIgNTQTPA
33	1063.4	99.8	2100	12	AAQ12072	T-PA with -ve char
34	1063.4	99.8	2170	8	AAN70248	Sequence encoding
35	1063.4	99.8	2170	16	AAQ86576	Human tissue plasm
36	1063.4	99.8	2249	11	AAQ05535	Plasmid pIgN delta
37	1063.4	99.8	2457	17	AAN60146	cDNA sequence enco
38	1063.4	99.8	2509	10	AAN90508	cDNA of human tiss
39	1063.4	99.8	2519	24	ABK52545	Human tissue plasm
40	1063.4	99.8	2519	24	ABL59544	Human tissue type
41	1063.4	99.8	2547	4	AAN30001	Sequence encoding
42	1063.4	99.8	7360	17	AAT15930	DHFR/intron (wTras
43	1062	99.7	1128	24	AAD40614	OmpA-K2S fusion pr
44	1062	99.7	1341	11	AAQ06762	Sequence encoding
45	1061.8	99.7	1068	10	AAN91122	Sequence of coding

ALIGNMENTS

RESULT 1
AAD40616
ID   AAD40616 standard; DNA; 1065 BP.
XX
AC   AAD40616;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Human kringle 2 serine protease (K2S) DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; enzyme; human; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1065
FT   /*tag= a
FT   /product= "Human K2S protein"
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.

```
Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCGGTCGGTCCGCACT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCGGTCGGTCCGCACT 660

Qy  661 GTGTGCCTTCCCCGGCCGGACCTGCAGCTGCCGCGACTGGACGACGAGGTGTGAGCTCCCGGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GTGTGCCTTCCCCGGCCGGACCTGCAGCTGCCGCGACTGGACGACGAGGTGTGAGCTCCCGGC 720

Qy  721 TACGGCAAGCATGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGTGAAGGAGGCTCATGTC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TACGGCAAGCATGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGTGAAGGAGGCTCATGTC 780

Qy  781 AGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTACAACAGTCACCGAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTACAACAGTCACCGAC 840

Qy  841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCGGGCGGCCCCAGGCAAACTTGCACGACGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCGGGCGGCCCCAGGCAAACTTGCACGACGCC 900

Qy  901 TGCCAGGGCGGATTCGGGGAGGAGGCCCCCTGGTGTGTCGGACGATGCGCCGATGACTTGGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TGCCAGGGCGGATTCGGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGCGCCGATGACTTGGTG 960

Qy  961 GGGCATCATCAGCTGGGGCTGGGCTGGCGTGTGGACAGAGAGGATGTCCCGGGTGTGTACACAAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 GGGCATCATCAGCTGGGGCTGGGCTGGCGTGTGGACAGAGAGGATGTCCCGGGTGTGTACACAAG 1020

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
             |||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
```

```
RESULT 2
AAN91120
ID   AAN91120 standard; DNA; 1068 BP.
XX
AC   AAN91120;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of coding region in plasmid pST112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key          Location/Qualifiers
FT   CDS          1..1065
FT                /*tag=  a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;  88EP-0112569.
XX
PR   03-AUG-1987;  87GB-0018298.
PR   26-OCT-1987;  87GB-0025052.
PR   13-NOV-1987;  87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94407.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for .
```

```
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25036.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
PS   Claim 12; Page 31; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is human K2S DNA.
XX
SQ   Sequence 1065 BP; 225 A; 314 C; 312 G; 214 T; 0 other;
```

```
Query Match             100.0%;  Score 1065;  DB 24;  Length 1065;
Best Local Similarity   100.0%;  Pred. No. 9.7e-237;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCGTGAATCCATGATCCTGATAGGCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCGTGAATCCATGATCCTGATAGGCAAGGTT 120

Qy  121 TACACAGCACAGAACCCAGTGCCCAGGCACTGCCCAGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TACACAGCACAGAACCCAGTGCCCAGGCACTGCCCAGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180

Qy  181 AATCCTGATAGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGCTGACGTGG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AATCCTGATAGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGCTGACGTGG 240

Qy  241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAG 300

Qy  301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCGACATGCGCCTCCCACCCCTGGCAGCTGCCATC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCGACATGCGCCTCCCACCCCTGGCAGCTGCCATC 360

Qy  361 TTTGCCAAGCACAGGAGGTGCGCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TTTGCCAAGCACAGGAGGTGCGCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420

Qy  421 TCCTGCTGGATTCTCTGCCGCGCCGCCACCTGCTTCCGGCCCCACCTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCTGCTGGATTCTCTGCCGCGCCGCCACCTGCTTCCGGCCCCACCTG 480

Qy  481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGTGCCCGGAGGAGGCAGAAATTTGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGTGCCCGGAGGAGGCAGAAATTTGAA 540

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
```

PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1068 BP; 226 A; 314 C; 313 G; 215 T; 0 other;

Query Match   100.0%;   Score 1065; DB 10; Length 1068;
Best Local Similarity   100.0%;   Pred. No. 9.7e-237;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTGGCACGCACAGC   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTGGCACGCACAGC   63

Qy   61 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 CTCACCGAGTCGGGTGCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   123

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGCGG   180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGCGG   183

Qy   181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGTGACGTGG   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGTGACGTGG   243

Qy   241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   303

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC   363

Qy   361 TTTGCCAAGCACAGGAGGAGTCGCCGGAGAGTCGGCCGGAGAGCGCATACTACTCATCAGC   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TTTGCCAAGCACAGGAGGAGTCGCCGGAGAGTCGGCCGGAGAGCGCATACTACTCATCAGC   423

Qy   421 TCCTGCTGGATTCTCTGCGCCCGCCACTGCTTCGCCCCCACCACCTG   480
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   424 TCCTGCTGGATTCTCTGCGCCCGCCACTGCTTCGCCCCCACCACCTG   483

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGTGGTCCCTGCGGGAGGAGCAGAAATTTGAA   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 ACGGTGATCTTGGGCAGAACATACCGGGTGTGGTCCCTGCGGGAGGAGCAGAAATTTGAA   543

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACAATTGCG   600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACAATTGCG   603

Qy   601 CTGCTGAGCTGAAATCGATTCGTCCCGCTGTCGCCCAGGAGAGCAGGGTGGTCCGCCACT   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604 CTGCTGAGCTGAAATCGATTCGTCCCGCTGTCGCCCAGGAGAGCAGCGTGGTCCGCCACT   663

Qy   661 GTGTGCCTTCCCCGGGGGACTGCAGCTGCGGCCGGAGTGTGAGCTCTCCGGC   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664 GTGTGCCTTCCCCGGGGGACTGCAGCTGCGGCCGGAGTGTGAGCTCTCCGGC   723

Qy   721 TACGGCAAGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAAGGAGCGGCTGAAGGAGGCTCATGTC   780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   724 TACGGCAAGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAAGGAGCGGCTGAAGGAGGCTCATGTC   783

Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGACAACAGTCACCGAC   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784 AGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGACAACAGTCACCGAC   843

Qy   841 AACATGCTGTGTCTGGAGACAACTCGGAGACGCGGGCCCCAGGCAAACTTGCACGACGCC   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844 AACATGCTGTGTCTGGAGACAACTCGGAGACGCGGGCCCCAGGCAAACTTGCACGACGCC   903

Qy   901 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGCCGATGACTTTGGTG   960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   904 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGCCGATGACTTTGGTG   963

Qy   961 GGCATCATCAGCTGGGGCCTGGGCTTCGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACAACAAAG   1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964 GGCATCATCAGCTGGGGCCTGGGCTTCGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACAACAAAG   1023

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA   1065
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA   1068

RESULT 3
AAN91135
ID   AAN91135 standard; DNA; 1170 BP.
XX
AC   AAN91135;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pmTTk.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1065
FT   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988; 88BP-0112569.
XX
PR   03-AUG-1987; 87GB-0018298.
PR   26-OCT-1987; 87GB-0025052.
PR   13-NOV-1987; 87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94417.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1170 BP; 249 A; 338 C; 346 G; 237 T; 0 other;

Query Match   100.0%;   Score 1065; DB 10; Length 1170;
Best Local Similarity   100.0%;   Pred. No. 9.9e-237;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTGGCACGCACAGC   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTGGCACGCACAGC   165

Qy   61 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   225

```
Qy  121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  226 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 285

Qy  181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  286 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG 345

Qy  241 GAGTACTGTGATGTGCCCTCTGCTCCACTGGGGCCTGAGACAGTAGCAGCCTCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  346 GAGTACTGTGATGTGCCCTCCTCTGCTCCACCTGCGGCCTGAGACAGTAGCAGCCTCAG 405

Qy  301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGCCTCCCACCCCTGGCAGGCTGCCATC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  406 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGCCTCCCACCCCTGGCAGGCTGCCATC 465

Qy  361 TTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCCGGTTCCGTTCGTGTGCGGAGAGTTTCCCC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  466 TTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCCGGTTCCGTTCGTGTGCGGAGAGTTTCCCC 525

Qy  421 TCCTGCTGGATTCTCTCGCCGCCCACTGTTCCCGCCACCTG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  526 TCCTGCTGGATTCTCTCGCCGCCCACTGCTTCCAGGACACCTG 585

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTTCCCTGCCGGGAGGAGCAGAAATTTGAA 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  586 ACGGTGATCTTGGGCAGAACATACCGGGTGGTTCCCTGCCGGGAGGAGCAGAAATTTGAA 645

Qy  541 GTGCGAAAAATACATTGTCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  646 GTGCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG 705

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  706 CTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 765

Qy  661 GTGTGCCCTTCCCCGGCGGACTTGACAGCTGCAGCTGGACCGAGTGTGAGCTCTCCGGC 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  766 GTGTGCCCTTCCCCGGCGGACTTGACAGCTGCAGCTGGACCGAGTGTGAGCTCTCCGGC 825

Qy  721 TACGGCAAGCAGCATGAGGGCCTTGCTGTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  826 TACGGCAAGCAGCATGAGGGCCTTGCTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 885

Qy  781 AGAGCTGTACCCATCCAGCCGCTGCACATCACAACAACATTTACTTAACAGAACAGTCACCGAC 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  886 AGAGCTGTACCCATCCAGCCGCTGCACATCACAACAACATTTACTTAACAGAACAGTCACCGAC 945

Qy  841 AACATGCTGTGTGCTGAGACACTCGAGCGGCGGCCCCAGGCAAACTTCGACGACGCC 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  946 AACATGCTGTGTGCTGAGACACTCGAGCGGCGGGCCCCCAGGCAAACTTCGACGACGCC 1005

Qy  901 TGGCCAGGGCGAATTCGGAGGAGGCCCCCCTCGTGTGTCTGAACGATGGCCCATGAGTTTCGGTG 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1006 TGGCCAGGGCGAATTCGGAGGAGGCCCCCCTCGTGTGTCTGAACGATGGCCCATGACTTTGGTG 1065

Qy  961 GGGCATCAGCTGGGGGCTGTGTGGACAGAAGGATGTCCCGGGGTGTGTACAAAAG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1066 GGGCATCAGCTGGGGGCCTGTGTGGACAGAAGGATGTCCCGGGGTGTGTACAAAAG 1125

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
        ||||||||||||||||||||||||||||||||||||||||||||
Db 1126 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1170
```

```
DT   03-OCT-2002   (updated)
DT   18-JUN-1990   (first entry)
XX
DE   Sequence of coding region in plasmid pTQKPA delta trp.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1065
FT                  /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;  88EP-0112569.
XX
PR   03-AUG-1987;  87GB-0018298.
PR   26-OCT-1987;  87GB-0025052.
PR   13-NOV-1987;  87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94409.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page 7; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1314 BP; 286 A; 386 C; 393 G; 249 T; 0 other;

Query Match      100.0%;  Score 1065;  DB 10;  Length 1314;
Best Local Similarity  100.0%;  Pred. No. 1e-236;
Matches 1065;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCGCTCACCGTGCACGCACAGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  250 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCGCTCACCGTGCACGCACAGC 309

Qy   61 CTCACCGAGTCGGGTGCGGTGCCTCCCGCCCGCCCCCGTGAATTCCATGATCCTGATAGGACAAGGTT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  310 CTCACCGAGTCGGGTGCGGTGCCTCCCGCCCGCCCCCGTGAATTCCATGATCCTGATAGGACAAGGTT 369

Qy  121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTAATTACTGCCGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  370 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTAATTACTGCCGG 429

Qy  181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  430 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG 489

Qy  241 GAGTACTGTGATGTGCCCTCCTCTGCTCCACCTGCGGCAGCAGTACTACAGCCTCAG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  490 GAGTACTGTGATGTGCCCTCCTCTGCTCCACCCCTGCGGCAGCAGTACTACAGCCTCAG 549

Qy  301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  550 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC 609

Qy  361 TTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  610 TTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 669
```

```
RESULT 4
AAN91123
ID   AAN91123  standard; DNA; 1314 BP.
XX
AC   AAN91123;
XX
DT   25-MAR-2003   (updated)
```

```
RESULT 5
AAN91133
ID   AAN91133 standard; DNA; 1419 BP.
XX
AC   AAN91133;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of coding region in plasmid pmTOk112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1065
FT                  /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94416.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;

Query Match             100.0%;  Score 1065;  DB 10;  Length 1419;
Best Local Similarity   100.0%;  Pred. No. 1e-236;
Matches  1065;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC    60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  355  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC   414

Qy   61  CTCACCGAGTCGGGTGCCTCCGTCCCGTGCCTCCCGTGCGAATTCCATGATCCTGATAGGCAAGGTT   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  415  CTCACCGAGTCGGGTGCCTCCGTCCCGTGCCTCCCGTGCGAATTCCATGATCCTGATAGGCAAGGTT   474

Qy  121  TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCTGGCCTGACAACATAATTACTGCCGG   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  475  TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCTGGCCTGACAACATAATTACTGCCGG   534

Qy  181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGTGAAGAACCGCAGGCTGACGTGG   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  535  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGTGAAGAACCGCAGGCTGACGTGG   594

Qy  241  GAGTACTGTGATGTGCCCCTGCTCACCTGCGGGCTCCACCTGCGGGCTCCAGCCTCAG   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  595  GAGTACTGTGATGTGCCCCTGCTCACCTGCGGGCTCCACCTGCGGGCTCCAGCCTCAG   654

Qy  301  TTTCGCATCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  655  TTTCGCATCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC   714

Qy  361  TTTGCCAAGCACAGGAGGTCGCCCGGAGGCGGTTCGCGGGGGCATACTACTCATCAGC   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  715  TTTGCCAAGCACAGGAGGTCGCCCGGAGGCGGTTCGCGGGGGCATACTACTCATCAGC   774

Qy  421  TCCTGCTGGATTCTCTGCGCCGCACTGCTTCCGCCCCACCACCTG   480
         |||||||||||||||||||||||||||||||||||||||||||||
Db  775  TCCTGCTGGATTCTCTGCGCCGCACTGCTTCCGCCCCACCACCTG   834

Qy  481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  835  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA   894

Qy  541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACATTGCG   600
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  895  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACATTGCG   954

Qy  601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGGCACT   660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  955  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGGCACT  1014

Qy  661  GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCCGGACTGTGAGCTGTGAGCTGTGAGCTCCCGGC   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1015  GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCCGGACTGTGAGCTGTGAGCTGTGAGCTCCCGGC  1074
```

```
Qy  421  TCCTGCTGGATTCTCTTGCCGGCCCACTGCTTCCTCAGAGAGGTTCCCGCCCACCACCTG   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  670  TCCTGCTGGATTCTCTTGCCGGCCCACTGCTTCCTCAGAGAGGTTCCCGCCCACCACCTG   729

Qy  481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  730  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA   789

Qy  541  GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACACAATGACATTGCG   600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  790  GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACACAATGACATTGCG   849

Qy  601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGGACT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  850  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGGACT   909

Qy  661  GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCCGGACTGTGAGCTCTCCGGC   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  910  GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCCGGACTGTGAGCTCTCCGGC   969

Qy  721  TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  970  TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  1029

Qy  781  AGACTGTACCCATCCAGCCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1030  AGACTGTACCCATCCAGCCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  1089

Qy  841  AACATGCTGTGTGGAGACACTCGGAGCGCGGGGCCCCTGTTGCACGACGCC   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1090  AACATGCTGTGTGGAGACACTCGGAGCGCGGGGCCCCTGTTGCACGACGCC   1149

Qy  901  TGCCAGGGCGATTCGGAGGAGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1150  TGCCAGGGCGATTCGGAGGAGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG  1209

Qy  961  GGCATCATCAGCTGGGGCCCTGGGCGTGTGGACAGAAGGATGTCCCGGTGTGTACACAAAG  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1210  GGCATCATCAGCTGGGGCCCTGGGCGTGTGGACAGAAGGATGTCCCGGTGTGTACACAAAG  1269

Qy 1021  GTTACCAACTACTAGACTGGATTCGTGACAACATGCGACCGTGA  1065
         ||||||||||||||||||||||||||||||||||||||||||||
Db 1270  GTTACCAACTACTAGACTGGATTCGTGACAACATGCGACCGTGA  1314
```

```
Qy  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGCTCATGTC  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1075 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGCTCATGTC 1134

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGAC  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1135 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGAC 1194

Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTGCACGACGCC  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1195 AACATGCTGTGTCTGGAGACACTCGGAGCACTCGGAGCGCGGGCGGCGGCAAACTGCACGACGCC 1254

Qy  901 TGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTGTGCTGAACGATGGCCGCATGACTTTGGTG  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1255 TGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTGTGCTGAACGATGGCCGCATGACTTTGGTG 1314

Qy  961 GGCATCATCAGCTGGGGCGCTGGGCCTGGCGTGTGCTGGACAGAGAAGATGTCCCCGGGTGTGTACACAAAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1315 GGCATCATCAGCTGGGGCGCTGGGCCTGGCGTGTGCTGGACAGAGAAGATGTCCCCGGGTGTGTACACAAAG 1374

Qy 1021 GTTACCAACTACTTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
        |||||||||||||||||||||||||||||||||||||||||||||
Db 1375 GTTACCAACTACTTAGACTGGATTCGTGACAACATGCGACCGTGA 1419
```

RESULT 6

AAT27588

ID   AAT27588 standard; DNA; 1680 BP.
XX
AC   AAT27588;
XX
DT   25-MAR-2003 (updated)
DT   06-AUG-1996 (first entry)
XX
DE   Novel plasminogen activator DNA.
XX
KW   Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
KW   fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
KW   protein engineering; kringle; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   sig_peptide   1..105
FT   /*tag= a
FT   mat_peptide   106..1677
FT   /*tag= b
XX
PN   US5504001-A.
XX
PD   02-APR-1996.
XX
PF   06-JUN-1994; 94US-0254485.
XX
PR   25-NOV-1987; 87US-0125629.
PR   28-JAN-1992; 92US-0827587.
PR   06-JUN-1994; 94US-0254485.
XX
PA   (ZYMO) ZYMOGENETICS INC.
XX
PI   Foster DC;
XX
DR   WPI; 1996-187699/19.
DR   P-PSDB; AAR96223.
XX
PT   Hybrid plasminogen activator comprises human tPA activator and
PT   N-terminal crosslinking domain from alpha2-plasmin inhibitor -
PT   useful to treat thrombosis and image blood clots
XX
PS   Example 3; Fig 11; 35pp; English.
XX
CC   A DNA construct (AAT27586) codes for a novel plasminogen
CC   activator (AAR96222) in which the kringle K1 domain (AAR96221)
CC   replaces the native K1 domain of tissue
CC   plasminogen activator (tPA) (AAR96220). It was obtd. by
CC   inserting a synthetic plasminogen K1 domain DNA (AAT27586)
CC   into tPA cDNA and can be expressed in transformed Escherichia
CC   coli RRI cells (FERM P-9272). Novel plasminogen activators are
CC   produced that show increased clot lysing specificity or plasma
CC   half-life.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 1680 BP; 387 A; 471 C; 479 G; 343 T; 0 other;

Query Match   100.0%;   Score 1065;   DB 17;   Length 1680;
Best Local Similarity 100.0%;   Pred. No. 1.1e-236;
Matches 1065;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGGCACAGC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  616 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGGCACAGC  675

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  676 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  735

Qy  121 TACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCAAACATAATTAATTACTGCCGG  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  736 TACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCAAACATAATTAATTACTGCCGG  795

Qy  181 AATCCTGATGGGGATGCCAAGCCCTGTGCCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  796 AATCCTGATGGGGATGCCAAGCCCTGTGCCCACGTGCTGAAGAACCGCAGGCTGACGTG  855

Qy  241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGCAGCCTCAG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  856 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGCAGCCTCAG  915

Qy  301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  916 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC  975

Qy  361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGGCCGGTTCTGTGCTGTGCGGGGGCATACTCATCAGC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  976 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGGCCGGTTCTGTGCTGTGCGGGGGCATACTCATCAGC 1035

Qy  421 TCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGGAGTTCCGCCGCCACCACCTG  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1036 TCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGGAGTTCCGCCGCCACCACCTG 1095

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCGGTGTCCCGGTGTCCCGGCAGAAATTTGAA  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1096 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCGGTGTCCCGGTGTCCCGGCAGAAATTTGAA 1155

Qy  541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1156 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1215

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTTGTGCCGTGCCGTGGTCCGCACT  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1216 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTTGTGCCGTGCCGTGGTCCGCACT 1275

Qy  661 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCAGCTGGACGTGGAGCTCTCCGGC  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1276 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCAGCTGGACGTGGAGCTCTCCGGC 1335

Qy  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGCTCATGTC  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1336 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGCTCATGTC 1395

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGAC  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1396 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGAC 1455

Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC  900
```

```
SQ   Sequence 1680 BP; 386 A; 471 C; 480 G; 343 T; 0 other;

     Query Match           100.0%;   Score 1065;   DB 17;   Length 1680;
     Best Local Similarity 100.0%;   Pred. No. 1.1e-236;
     Matches 1065;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGCACGCACAGC   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   616 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGCACGCACAGC  675

Qy    61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   676 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  735

Qy   121 TACACAGCACAGACCCCAGTGCCAGGCACTGGGCTCGGGCAAACATAATTACTGCCGG  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   736 TACACAGCACAGACCCCAGTGCCAGGCACTGGGCTCGGGCAAACATAATTACTGCCGG  795

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   796 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  855

Qy   241 GAGTACTGTGATGTGCCCTGCTCCACCTGGCGGCCTGAGACAGTACAGCCAGCCTCAG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   856 GAGTACTGTGATGTGCCCTGCTCCACCTGGCGGCCTGAGACAGTACAGCCAGCCTCAG  915

Qy   301 TTTCGCATCAAAGAGGAGGGCTTCTCGCCGCACATGCCGCCTCCCCGGCAGGCTGCCATC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   916 TTTCGCATCAAAGAGGAGGGCTTCTCGCCGCACATGCCGCCTCCCCGGCAGGCTGCCATC  975

Qy   361 TTTGCCAAGCACAGCAGGAGTCGGCCGGTTCGTGTGCGGGAGGCCATACTCATCAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   976 TTTGCCAAGCACAGCAGGAGTCGGCCGGTTCGTGTGCGGGAGGCCATACTCATCAGC 1035

Qy   421 TCCTGCTGATTCTCTCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCCACCACCTG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1036 TCCTGCTGATTCTCTCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCCACCACCTG 1095

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1096 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA 1155

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACGACAATGACATTGCCG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1156 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACGACAATGACATTGCCG 1215

Qy   601 CTGCTGCAGCTGAAATCGGATTCTCCCGCGTGTCCCAGGAGAGCAGCGTGGTCCGCACT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1216 CTGCTGCAGCTGAAATCGGATTCTCCCGCGTGTCCCAGGAGAGCAGCGTGGTCCGCACT 1275

Qy   661 GTGTGCCTTCCCCGGGCCAGCTGCAGCTGACGGACTGACGAGTGTGAGCTCTCCGGC  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1276 GTGTGCCTTCCCCGGGCCAGCTGCAGCTGACGGACTGACGAGTGTGAGCTCTCCGGC 1335

Qy   721 TACGGCAAGCAAGCATGAGGCCTTGTCTCCTTTCTCCTTTCTATTCGGAGCGGCTCATGTC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1336 TACGGCAAGCAAGCATGAGGCCTTGTCTCCTTTCTCCTTTCTATTCGGAGCGGCTCATGTC 1395

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1396 AGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1455

Qy   841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCCCCAGCAAACTTGCACGACGCGCC  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1456 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCCCCAGCAAACTTGCACGACGCGCC 1515

Qy   901 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCATGACTTTGGTG  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1516 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCATGACTTTGGTG 1575

Qy   961 GGCATCATCAGCTGGGGCTGGGCCTGGGCTGTGCTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1576 GGCATCATCAGCTGGGGCTGGGCCTGGGCTGTGCTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1635

Qy  1021 GTTACCAACTAACTACCTAGACTCGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1636 GTTACCAACTAACTACCTAGACTCGATTCGTGACAACATGCCGACCGTGA 1680
```

```
RESULT 7
AAT27587
ID   AAT27587 standard; DNA; 1680 BP.
XX
AC   AAT27587;
XX
DT   25-MAR-2003  (updated)
DT   06-AUG-1996  (first entry)
XX
DE   Novel plasminogen activator DNA.
XX
KW   Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
KW   fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
KW   protein engineering; kringle; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   sig_peptide    1..105
FT                  /*tag= a
FT   mat_peptide    106..1677
FT                  /*tag= b
XX
PN   US5504001-A.
XX
PD   02-APR-1996.
XX
PF   06-JUN-1994;  94US-0254485.
XX
PR   25-NOV-1987;  87US-0125629.
PR   28-JAN-1992;  92US-0827587.
PR   06-JUN-1994;  94US-0254485.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Foster DC;
XX
DR   WPI; 1996-187699/19.
DR   P-PSDB; AAR96222.
XX
PT   Hybrid plasminogen activator comprises human tPA activator and
PT   N-terminal crosslinking domain from alpha2-plasmin inhibitor -
PT   useful to treat thrombosis and image blood clots
XX
PS   Example 3; Fig 10; 35pp; English.
XX
CC   A DNA construct (AAT27587) codes for a novel plasminogen
CC   activator (AAR96223) in which the kringle K1 domain of plasminogen
CC   (see also AAR96221) is mutated to substitute asparagine for aspartic
CC   acid at position 5, and replaces the native K1 domain of tissue
CC   plasminogen activator (tPA) (AAR96220). It was obtd. by mutagenesis
CC   of the sequence in vector Zem99-8000 (see also AAT27587). The novel
CC   protein can be expressed in Escherichia coli RRI/Zem99-8010 cells
CC   (FERM P-9315).  Novel plasminogen activators are produced that show
CC   increased clot lysing specificity or plasma half-life.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
```

```
Db   745  TACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG   804

Qy   181  AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGCACGTGG   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805  AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGCACGTGG   864

Qy   241  GAGTACTGTGATGTGCCTCCTGCTCCACCTGGCGGCCTGAGACAGTACAGCCAGCCTCAG   300
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   865  GAGTACTGTGATGTGCCTGAGCTGCTCCACCTGGCGGCCTGAGACAGTACAGCCAGCCTCAG   924

Qy   301  TTTCGCATCAAAGGAGGGCTCTTGCCGGCATGCCTCCACCCCTGGCAGGCTGCCATC   360
          ||||||||||||||||||||||| |||||||| |||||||||||||||||||||||||
Db   925  TTTCGCATCAAAGGAGGGCTCTTCGCCGGCATCCCTCCACCCCTGGCAGGCTGCCATC   984

Qy   361  TTTGCCAAGCACAGGAGGTCGCCCGGAGACGGTTCCTGTGCGGGGCATACTCATCAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985  TTTGCCAAGCACAGGAGGTCGCCCGGAGACGGTTCCTGTGCGGGGCATACTCATCAGC   1044

Qy   421  TCCTGCTGGATTCTCTGCGGCCCACTGCTTCCAGGAGAGTTCCGCCCACCACCTG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1045  TCCTGCTGGATTCTCTGCGGCCCACTGCTTCCAGGAGAGTTCCGCCCACCACCTG   1104

Qy   481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGGAGGAGCAGAAATTTGAA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1105  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGGAGGAGCAGAAATTTGAA   1164

Qy   541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAATCGATTGACAATGACATTGCG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1165  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAATCGATTGACAATGACATTGCG   1224

Qy   601  CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCGTGGTCCGCACT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1225  CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCGTGGTCCGCACT   1284

Qy   661  GTGTGGCCTTCCCCGGCGGACCTGCAGCTGCCGGACCTGCCGGAGTGTGAGCTCTCCGGC   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1285  GTGTGGCCTTCCCCGGCGGACCTGCAGCTGCCGGACCTGCCGGAGTGTGAGCTCTCCGGC   1344

Qy   721  TACGGCAAGCAGCATGAGGAGGCCTTGTCTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1345  TACGGCAAGCAGCATGAGGAGGCCTTGTCTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   1404

Qy   781  AGACTGTACCCATCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1405  AGACTGTACCCATCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   1464

Qy   841  AACATGCTGTGTCTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1465  AACATGCTGTGTCTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC   1524

Qy   901  TGCCAGGCGGATTCGGGAGGCGCCCCTGGTGTGTCTGAACGATGCCGATGACTTGGTG   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  1525  TGCCAGGCGGATTCGGGAGGCGCCCCTGGTGTGTCTGAACGATGCCGATGACTTTGGTG   1584

Qy   961  GGCATCATCAGCTGGGGCTGGGCTGTGGACGAGAGGATGTCCCGGGTGTGTACACAAAG   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1585  GGCATCATCAGCTGGGGCTGGGCTGTGGACGAGAGGATGTCCCGGGTGTGTACACAAAG   1644

Qy  1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA   1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1645  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA   1689
```

```
Qy  1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA   1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1636  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA   1680
```

RESULT 8
AAN82178
ID   AAN82178 standard; DNA; 1689 BP.
XX
AC   AAN82178;
XX
DT   25-MAR-2003 (updated)
DT   10-MAR-2003 (updated)
DT   03-NOV-1990 (first entry)
XX
DE   Tissue plasminogen activator with S-119 substd for M.
XX
KW   Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.
XX
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1689
FT   /tag= a
FT   /product=modified tPA
XX
PN   JP63230083-A.
XX
PD   26-SEP-1988.
XX
PF   20-MAR-1987; 87JP-0064339.
XX
PR   20-MAR-1987; 87JP-0064339.
XX
PA   (EISA) EISAI CO LTD.
XX
DR   WPI; 1988-311961/44.
DR   P-PSDB; AAP82581.
XX
PT   Modified tissue plasminogen activator - having glycine-183 and serine-186
PT   residues sustd. with serine and threonine.
XX
PS   Disclosure; Page 7; 16pp; Japanese.
XX
CC   One N-glycosylation site, i.e. NSS (117-119) is substituted with
CC   NSM and the N-glycosylation is removed.
CC   Plasmid encoding the modified tPA is 99-6300 and its transformant is
CC   E.coli RR1-Zem 99-6300 (FERM P-9127).
CC   This modified tPA, used to treat thrombosis, is of high quality and
CC   has a longer half life period in blood.
CC   See also AAN82177-N82179.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

Query Match   100.0%; Score 1065; DB 9; Length 1689;
Best Local Similarity   100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCACGCACAGC   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCACGCACAGC   684

Qy    61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTGGAATTCCATGATCCTGATAGGCAAGGTT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTGGAATTCCATGATCCTGATAGGCAAGGTT   744

Qy   121  TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 9
AAQ01358
ID   AAQ01358 standard; DNA; 1780 BP.
XX
AC   AAQ01358;
XX
DT   25-MAR-2003 (updated)
DT   08-SEP-1990 (first entry)
XX
DE   Sequence encoding wild type tissue plasminogen activator (t-PA).

XX
KW   Wild type tissue plasminogen activator (t-PA); infarction treatment;
KW   thrombosis treatment; embolism treatment.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   1..123
FT   /*tag= a
FT   /note="synthetic 5' adaptor"
FT   misc_feature   1738..1780
FT   /*tag= b
FT   /note="synthetic 3' adaptor"
XX
PN   EP351246-A.
XX
PD   17-JAN-1990.
XX
PF   14-JUL-1989; 89EP-0307194.
XX
PR   15-JUL-1988; 88DK-0003952.
XX
PA   (NOVO) NOVO-NORDISK AS.
XX
PI   Petersen LC, Boel E;
XX
DR   WPI; 1990-016567/03.
DR   P-PSDB; AAR04699.
XX
PT   New tissue plasminogen activator (t-PA) analogue -
PT   with higher fibrin selectivity than native t-pa. useful for
PT   treating infarction, thrombosis and embolism
XX
PS   Disclosure; Fig 5A-D; 24pp; English.
XX
CC   It is modified in the patent by replacing one or more codons specifying
CC   a positively charged amino acid. The resulting analogues are inserted
CC   into a replicable expression vector which is used to transform or
CC   transfect a host cell which is grown to express a t-PA analogue. The
CC   t-PA analogue is useful for the treatment of diseases or disorders
CC   associated with the formation of thrombi in blood vessels e.g.
CC   infarctions, thrombosis and embolism. The analogue, in the 1-chain form,
CC   exhibits the properties of a proenzyme; however on plasmin-catalysed
CC   cleavage of the one-chain form, the activity of the 2-chain form is fully
CC   retained. Compared to native t-PA, it has a higher fibrin selectivity,
CC   this results in a fibrinolytic agent with a higher fibrin selectivity,
CC   as the fibrinogenolytic activity induced by the analogue is reduced
CC   relative to that induced by native 1-chain t-PA.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;

Query Match   100.0%; Score 1065; DB 11; Length 1780;
Best Local Similarity   100.0%; Pred. No. 1.1e-236;
Matches 1065   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCAACCGGTGGCACGCACAGC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   641 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCAACCGGTGGCACGCACAGC 700

Qy   61 CTCACCGGTCGGGTGCCTCCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   701 CTCACCGGTCGGGTGCCTCCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 760

Qy   121 TACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   761 TACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 820

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGACCCGGCTGCTGACGTGG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   821 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGACCCGGCTGCTGACGTGG 880

Qy   241 GAGTACTGTGATGTGATGTGCCCTCCTGCTCCACCTGCCGGCCGGCCTGAGACAGTACAGCCGCCTCAG 300

Db   881 GAGTACTGTGATGTGATGTGCCCTCCTGCTCCACCTGCCGGCCGGCCTGAGACAGTACAGCCGCCTCAG 940
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 TTTGCCATCAAAGAGGAGGGCTCTTGCCAGACACGCCATCCCTGGCAGGCTGCCATC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   941 TTTGCCATCAAAGAGGAGGGCTCTTGCCAGACACGCCATCCCTGGCAGGCTGCCATC 1000

Qy   361 TTTGCCAAGCACAGAGGTCGCCGGAGACGGGTTCGTGCGGGGGCATACTACTCATCAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1001 TTTGCCAAGCACAGAGGTCGCCGGAGACGGGTTCGTGCGGGGGCATACTACTCATCAGC 1060

Qy   421 TCCTGCTGATTCTCTGCTGCTTCCAGGACAGAGGTTTCCGCCCACCACCTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1061 TCCTGCTGATTCTCTGCTGCTTCCAGGACAGAGGTTTCCGCCCACCACCTG 1120

Qy   481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1121 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 1180

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1181 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1240

Qy   601 CTGCTGCAGCTGAAATCGAATTCGTCCCGTGTGCCCAGGAGGAGCAGCGGTGGTCCGCACT 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1241 CTGCTGCAGCTGAAATCGAATTCGTCCCGTGTGCCCAGGAGGAGCAGCGGTGGTCCGCACT 1300

Qy   661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGACTGGACGGAGTGTGAGCTCTCCGGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1301 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGACTGGACGGAGTGTGAGCTCTCCGGC 1360

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1361 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1420

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1421 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1480

Qy   841 AACATGCTGTGTCTGGAGACACTCGAGCGGCGGCCCAGCCAAACTTGCACGACGCGCC 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1481 AACATGCTGTGTCTGGAGACACTCGAGCGGCGGCCCAGCCAAACTTGCACGACGCGCC 1540

Qy   901 TGCCAGGCGGATTCGGAGGAGCCCCCCTGGTGTGTGACGGCCGCATGAGACTTGGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1541 TGCCAGGCGGATTCGGAGGAGCCCCCTGGTGTGTGACGGCCGCATGAGACTTGGTG 1600

Qy   961 GGCATCATCAGCTGGGGCTGGGCGTGTGACAGAGGATGTCCCGGGGTGTACACAAAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1601 GGCATCATCAGCTGGGGCTGGGCGTGTGACAGAGGATGTCCCGGGGTGTGTACACAAAG 1660

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCGTGA 1065
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   1661 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCGTGA 1705

RESULT 10
AAN91119
ID   AAN91119 standard; DNA; 2100 BP.
XX
AC   AAN91119;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of native tPA in plasmid pSTi12.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX

```
FH          Key             Location/Qualifiers
FT          CDS             25..1710
FT                          /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;    88EP-0112569.
XX
PR   03-AUG-1987;    87GB-0018298.
PR   26-OCT-1987;    87GB-0025052.
PR   13-NOV-1987;    87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94406.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Example 29; Fig 21; 68pp; English.
XX
CC   cDNA sequence of native tPA gene is excised from plasmid pSTI12, and
CC   digested with BamHI and SalI to form plasmid pSTI18.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match             100.0%;  Score 1065;  DB 10;  Length 2100;
Best Local Similarity   100.0%;  Pred. No. 1.1e-236;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTGGCACGCACAGC   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTGGCACGCACAGC  708

Qy         61 CTCACCGAGTCGGGTGCCCTCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 CTCACCGAGTCGGGTGCCCTCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT  768

Qy        121 TACAACGACACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACTGGGCAAACATAATTACTGCCGG  180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 TACAACGACACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACTGGGCAAACATAATTACTGCCGG  828

Qy        181 AATCCTGATGGGAGCCAAGCCCTGGTGCCACGTGCCAGCCCTGGTGCCACGTGGCTGGG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 AATCCTGATGGGAGCCAAGCCCTGGTGCCACGTGCCAGCCCTGGTGCCACGTGGCTGGG  888

Qy        241 GAGTACTGTGATGTGCCCTCCTGTCCACCTGGCCGGGCCCTGAGACAGTACAGCCAGCCTCAG  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GAGTACTGTGATGTGCCCTCCTGTCCACCTGGCCGGGCCCTGAGACAGTACAGCCAGCCTCAG  948

Qy        301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTTCGCCGACATCGCCTGCCATC  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTTCGCCGACATCGCCTGCCATC 1008

Qy        361 TTTGCCAAGCACAGGAGGTCGCCCGGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 TTTGCCAAGCACAGGAGGTCGCCCGGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1068

Qy        421 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCTG  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCTG 1128

Qy        481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 1188
```

```
Qy        541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1248

Qy        601 CTGCTGCAGCTGAAAATCGGATTCGTCCGCTGTGCCCGTGGTGCCGCTCCGCACT  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1249 CTGCTGCAGCTGAAAATCGGATTCGTCCGCTGTGCCCGTGGTGCCGCTCCGCACT 1308

Qy        661 GTGTGCCTTCCCCCGGCCGACCTGCAGCTGCCCGACTGGACGGAGTGTGAGCTCTCCGGC  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 GTGTGCCTTCCCCCGGCCGACCTGCAGCTGCCCGACTGGACGGAGTGTGAGCTCTCCGGC 1368

Qy        721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGCGGCTGAAGGAGGCTCATGTC  780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGCGGCTGAAGGAGGCTCATGTC 1428

Qy        781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1429 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1488

Qy        841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC  900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1489 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1548

Qy        901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGCGGCCGATGACTTTGGTG  960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1549 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGCGGCCGATGACTTTGGTG 1608

Qy        961 GGCATCATCAGCTGGGGCTGGGCCTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1609 GGCATCATCAGCTGGGGCTGGGCCTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1668

Qy       1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1713
```

```
RESULT 11
AAQ05532
ID   AAQ05532 standard; DNA; 2100 BP.
XX
AC   AAQ05532;
XX
DT   25-MAR-2003 (updated)
DT   11-DEC-1990 (first entry)
XX
DE   Plasmid pSTI12 encoding novel N-terminal for tissue plasminogen
DE   activator (tPA).
XX
KW   Fibrin; clotting; thrombolytic; vascular disease; stroke;
KW   myocardial infarction; heart attack; pulmonary embolism; ds;
XX
OS   Homo sapiens.
XX
FH          Key             Location/Qualifiers
FT          CDS             25..1710
FT                          /*tag= a
FT          mat_peptide     130..1710
FT                          /*tag= b
XX
PN   EP379890-A.
XX
PD   01-AUG-1990.
XX
PF   10-JAN-1990;    90EP-0100457.
XX
PR   23-JAN-1989;    89GB-0001422.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Satoh S, Suzuki S, Otsuka K, Kusunoki C;
XX
```

```
DR   WPI; 1990-232757/31.
DR   P-PSDB; AAR06237.
XX
PT   New tissue plasminogen activator - having N-terminal peptide of
PT   plasminogen linked to tissue plasminogen activator for increased
PT   stability in vivo.
XX
PS   Disclosure; Page 7; ?pp; English.
XX
CC   tPA with N-terminal peptide of plasminogen linked is more stable in
CC   vivo than the native form. It is useful as a thrombolytic agent in
CC   the treatment of vascular diseases eg myocardial infarction,
CC   pulmonary embolism etc.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match             100.0%;  Score 1065;  DB 11;  Length 2100;
Best Local Similarity   100.0%;  Pred. No. 1.1e-236;
Matches 1065;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 708

Qy     61 CTCACCGAGTCGGTGCCCTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 CTCACCGAGTCGGTGCCCTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 768

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 828

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 888

Qy    241 GAGTACTGTGATGTGCCCTCTGCTCCACCTGGGGCTGGGGCCTGAGACAGTACAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 GAGTACTGTGATGTGCCCTCTGCTCCACCTGGGGCTGGGGCCTGAGACAGTACAGCCAGCCTCAG 948

Qy    301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGGCCTCCACCCTGCAGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    949 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGGCCTCCACCCTGCAGCTGCCATC 1008

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1068

Qy    421 TCCTGTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1069 TCCTGTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACTG 1128

Qy    481 ACGGTGATCTTGGGCAGAACATACCCGGGTGGTCCCTGCGAGGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 ACGGTGATCTTGGGCAGAACATACCCGGGTGGTCCCTGCGAGGAGGAGCAGAAATTTGAA 1188

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 GTGGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1248

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCGTGTGCCCGACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1249 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCGTGTGCCCGACT 1308

Qy    661 GTGTGCCTTCCCCCGGGCGACCTGCAGCTGACCGGAGTGAGCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1309 GTGTGCCTTCCCCCGGGCGACCTGCAGCTGACCGGAGTGAGCTCCGGC 1368

Qy    721 TACGGCAAGCAATGAGGGCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369 TACGGCAAGCAATGAGGGCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC 1428
```

```
Qy    781 AGACTGTACCCATCCAGCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429 AGACTGTACCCATCCAGCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1488

Qy    841 AACATGCTGTGTCGTGGAGACACTCGGACGCGGCGCGGCGGCGGCCCCAGGCAAACTTGCACGACGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489 AACATGCTGTGTCGTGGAGACACTCGGACGCGGCGCGGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1548

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCGAACGATGGCCGCATGAGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCGAACGATGGCCGCATGAGACTTTGGTG 1608

Qy    961 GGCATCATCAGCTGGGGCCTGGGCTGTGGCACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1609 GGCATCATCAGCTGGGGCCTGGGCTGTGGCACAGAAGGATGTCCCGGGTGTGTACACAAAG 1668

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCTGACCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCTGACCGACCGTGA 1713


RESULT 12
AAQ87370
ID   AAQ87370 standard; DNA; 2162 BP.
XX
AC   AAQ87370;
XX
DT   25-MAR-2003 (updated)
DT   19-SEP-1995 (first entry)
XX
DE   Human tPA clone.
XX
KW   Intron; recombination; combinatorial gene; trans-splicing;
KW   gene therapy; polymerase chain reaction; PCR; primer; amplification;
KW   tissue plasminogen activator; tPA; plasmid TPA-KS+; thrombolytic;
KW   ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     82..334
FT                   /*tag= a
FT                   /note="signal sequence and finger-like domain"
FT   CDS             335..447
FT                   /*tag= b
FT                   /product= EGF-like domain
FT   CDS             448..714
FT                   /*tag= c
FT                   /product= Kringle-1 domain
FT   CDS             715..972
FT                   /*tag= d
FT                   /product= Kringle-2 domain
FT   CDS             973..2162
FT                   /*tag= e
FT                   /product= catalytic domain
XX
PN   WO9507351-A1.
XX
PD   16-MAR-1995.
XX
PF   12-SEP-1994; 94WO-US10146.
XX
PR   10-SEP-1993; 93US-0119512.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Jarrell KA;
XX
DR   WPI; 1995-123425/16.
XX
PT   New intron-mediated recombinant techniques - used for the
PT   generation and selection of novel genes and gene prods. for use
PT   in therapy
```

```
XX
PS   Example 4; Page 59-60; 87pp; English.
XX
CC   A cDNA clone of human tissue plasminogen activator (tPA) was
CC   amplified by PCR using the primers given in AAQ87368-69.  The
CC   amplified tPA DNA (AAQ87370) was ligated into vector KS+ to
CC   obtain plasmid tPA-KS+.  The construct was used in combinatorial
CC   methods involving RNA splicing-mediated shuffling of tPA domains
CC   in plasmid pINVl (AAQ87347) to generate novel tPAs having
CC   improved thrombolytic properties.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match             100.0%;  Score 1065;  DB 16;  Length 2162;
Best Local Similarity   100.0%;  Pred. No. 1.1e-236;
Matches 1065;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTGGCACGCACAGC 765

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGTCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGTCAAGGTT 825

Qy    121 TACCACGACCAGAACCCCAGTGCCAGCACTGGGGCTGGCCAGCACATAATTACTGCCGG 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TACCACGACCAGAACCCCAGTGCCCAGCACTGGGGCTGGCCAAACATAATTACTGCCGG 885

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAAACCGCAGGCTGACGTGG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAAACCGCAGGCTGACGTGG 945

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTACAGCAGCCAGCCTCAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTACAGCAGCCAGCCTCAG 1005

Qy    301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATGCGCCAGCAGCGGCTGCCATC 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCCACATGCGCCAGCAGCGGCTGCCATC 1065

Qy    361 TTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 TTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 1125

Qy    421 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCGCGCCCCCACCACCTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCGCGCCCCCACCACCTG 1185

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGCGAGGAGGAGCAGAAATTTGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGCGAGGAGGAGCAGAAATTTGAA 1245

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1305

Qy    601 CTGCTGCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTCCGCACT 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 CTGCTGCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTCCGCACT 1365

Qy    661 GTGTGGCCTTCCCCCGGCGGGCCGGACTGGCCAGCTGCGACTGGACTGAGCTGTGAGCTTCCCGGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 GTGTGGCCTTCCCCCGGCGGGCCGGACTGGCCAGCTGCGACTGGACTGAGCTGTGAGCTTCCCGGC 1425

Qy    721 TACGGCAAGCAATGAGGCCTTGTCCTTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 TACGGCAAGCAATGAGGCCTTGTCCTTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 1485

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGACAACAGTCACCGAC 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGACAACAGTCACCGAC 1545

Qy    841 AACATGCTGTGTGCTGAGACACTCGGAGCGCGGCGGGGGGCCCAGGCAAACTTGCACGACGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 AACATGCTGTGTGCTGAGACACTCGGAGCGCGGCGGGGGGCCCAGGCAAACTTGCACGACGCC 1605

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTCTGAACGACGGCCGCATGACTTTGGTG 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTCTGAACGACGGCCGCATGACTTTGGTG 1665

Qy    961 GGCATCATCAGCTGGGCGCTGGGGCTGTGGACAGAAGGATGTCCCGGGTGTACACAAAG 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 GGCATCATCAGCTGGGCGCTGGGGCTGTGGACAGAAGGATGTCCCGGGTGTACACAAAG 1725

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
            |||||||||||||||||||||||||||||||||||||||||||||
Db   1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1770
```

CC   The present sequence represents the human tissue plasminogen activator
CC   gene. It was used to construct plasmid TPA-KS+, which is used in the
CC   course of the invention. The specification describes a purified
CC   reverse-splicing intron which comprises a segment comprising a
CC   5'-part of a group II intron, including an exon binding site not
CC   naturally present in the intron and a second segment comprising a
CC   3'-part of a group II intron, including a domain V motif, a branch
CC   site acceptor, and a nucleophilic group for transesterifying a
CC   phosphodiester bond of an RNA. Together the two segments form an
CC   autocatalytic Y-branched intron which catalyses integration of at least
CC   the first segment into substrate RNA by a reverse splicing reaction
CC   The reverse-splicing introns are used, by specific cleavage and ligation
CC   of discontinuous nucleic acid, to generate new genes and gene products,
CC   e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC   manipulation, e.g. replacements for restriction enzymes) or
CC   immunologically active or signal-transducing proteins such as antibody
CC   and enzyme libraries.
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match   100.0%;   Score 1065; DB 19;   Length 2162;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy        1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      706 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 765

Qy       61 CTCACCGAGTCGGGTGCCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      766 CTCACCGAGTCGGGTGCCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 825

Qy      121 TACACAGCACAGAACCCCAGTGCCCAGGACTGGGCCTGGCCTGGCAAACATAATTACTGCCGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      826 TACACAGCACAGAACCCCAGTGCCCAGGACTGGGCCTGGCCTGGCAAACATAATTACTGCCGG 885

Qy      181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      886 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG 945

Qy      241 GAGTACTGTGATGTGCCCCTGCCTCCACCTGGGCCGGCCTGAGACAGTACAGCCAGCCTCAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      946 GAGTACTGTGATGTGCCCCTGCCTCCACCTGGGCCGGCCTGAGACAGTACAGCCAGCCTCAG 1005

Qy      301 TTTCGCATCAAAGGAGAGGGCTCTTCGCCCGACATGCCATCGCCTCCACCCCTGCCAGGCTGCCATC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1006 TTTCGCATCAAAGGAGAGGGCTCTTCGCCCGACATGCCATCGCCTCCACCCCTGCCAGGCTGCCATC 1065

Qy      361 TTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1066 TTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1125

Qy      421 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCGGCGAGGAGGTTTCCGCCCACCTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1126 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCGGCGAGGAGGTTTCCGCCCACCTG 1185

Qy      481 ACGGTGATCTTGGCAGAACATACCGGGGTGGTCCCTGGCGGAGGAGGGAGAAATTTGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1186 ACGGTGATCTTGGCAGAACATACCGGGGTGGTCCCTGGCGGAGGAGGGAGAAATTTGAA 1245

Qy      541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGACATTGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGACATTGCG 1305

Qy      601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGTCCGGTGTGCCCAGGAGAGCAGCGGTGTCCGCACT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1306 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGTCCGGTGTGCCCAGGAGAGCAGCGGTGTCCGCACT 1365

Qy      661 GTGTGCCTTCCCCCGGCGACTTGCGCAGCTGCCGACTGACGGAGTGTGAGCTCTCCGGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1366 GTGTGCCTTCCCCCGGCGACTTGCGCAGCTGCCGACTGACGGAGTGTGAGCTCTCCGGC 1425

Qy      721 TACGGCAAGCAATGAGGCGTGGCCTTGCTCCTTTCTATTCGAGCGGCGGCTGAAGGAGGCTCATGTC 780
```

(continued with additional alignment blocks)

```
Db     1426 TACGGCAAGCATGAGGCATGAGAGCCTGTGTCTTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 1485

Qy      781 AGACTGTACCCATCCAGCGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1486 AGACTGTACCCATCCAGCGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1545

Qy      841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCCCGGGCCCCAGGCAAACTTGCACGACGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1546 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCCCGGGCCCCAGGCAAACTTGCACGACGCC 1605

Qy      901 TGCCAGGGGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1606 TGCCAGGGGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACTTTGGTG 1665

Qy      961 GGCATCATCAGCTGGGGCCTGGGCTTGTCTGGACAGAAGGATGTCCCGGGTGTCTACACAAAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1666 GGCATCATCAGCTGGGGCCTGGGCTTGTCTGGACAGAAGGATGTCCCGGGTGTCTACACAAAG 1725

Qy     1021 GTTACCAACTACCTAGACTGGATTCGTGACGACTGGAACAATGCGACCGACCGTGA 1065
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGACCGTGA 1770
```

RESULT 14
AAN60659
ID   AAN60659 standard; DNA; 2547 BP.
XX
AC   AAN60659;
XX
DT   25-MAR-2003 (updated)
DT   22-JUL-1991 (first entry)
XX
DE   Sequence encoding human pre-tissue plasminogen activator (pre-t-PA).
XX
KW   Plasminogen conversion; vascular disease therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   87..191
FT   /*tag= a
FT   192..1775
FT   /*tag= b
FT   mat_peptide
XX
PN   GB2173804-A.
XX
PD   22-OCT-1986.
XX
PF   21-APR-1986; 86GB-0609683.
XX
PR   22-APR-1985; 85US-0725468.
PR   01-APR-1986; 86US-0846697.
PR   22-APR-1985; 85US-0725468.
PR   01-APR-1986; 86US-0846697.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Heyneker HL, Vehar GA;
XX
DR   WPI; 1986-280715/43.
DR   P-PSDB; AAP60790.
XX
PT   New mutant forms of human tissue plasminogen activator- having
PT   higher specific activity and resistance to conversion to two chain
PT   form
XX
PS   Example; Fig 2; 34pp; English.
XX
CC   The patentors claim a novel recombinant human t-PA which is
CC   resistant to specific enzymatic cleavage because it is stabilised by
CC   site-directed mutagenesis at a 2-chain cleavage site. Partic. the
CC   natural Arg at position 275 is replaced by Gly or Glu, or Ile at

CC   position 276.is replaced (275 and 276 refer to the posn. of the AAs
CC   .in the mature protein; i.e. AAs 310 and 311 of AAP60790. Also new are
CC   (1) DNA sequences; (2) expression vectors; and (3) microorganisms
CC   and cell cultures transformed with these vectors .
CC   (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;

Query Match   100.0%; Score 1065; DB 7; Length 2547;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   711 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC 770

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   771 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 830

Qy   121 TACCACGACAGAACCCCAGTGCCCAGGCACTGGGCTGGGCCTGGCACAACATAATTACTGCGG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   831 TACCACGACAGAACCCCAGTGCCCAGGCACTGGGCTGGGCCTGGCACAACATAATTACTGCGG 890

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   891 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 950

Qy   241 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGGCCTGAGACCAGTACAGCCAGCCTCAG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   951 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGGCCTGAGACCAGTACAGCCAGCCTCAG 1010

Qy   301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGCCGACATGCCGCAGCTGCCATC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGCCGACATGCCGCAGCTGCCATC 1070

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 1130

Qy   421 TCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 TCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 1190

Qy   481 ACGGTGATCTTGGCAGAACATACCGGGTGCTGGTCCTGCGAGGAGGCAGAAAATTTGAA 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 ACGGTGATCTTGGCAGAACATACCGGGTGCTGGTCCTGCGAGGAGGCAGAAAATTTGAA 1250

Qy   541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGCAACTTACGACAATGACATTGCG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGCAACTTACGACAATGACATTGCG 1310

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCGCTGTGTCCCGTGTGCCCGCACT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 CTGCTGCGACTGAAATCGGATTCGTCCCGCTGTGCCCGTGTGCCCGGCAGCAGCGCAGCGCAGCGTGGTTCCGCACT 1370

Qy   661 GTGTGGCCTTCCCCGGCGACTTGACTCGGACTGACGGAGTGTGAGCGCTCCGGC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 GTGTGGCCTTCCCCGGCGACTTGACTCGGACTGACGGAGTGTGAGCGCTCCGGC 1430

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1490

Qy   781 AGACTGTACCCATCCAGCCGCTGAAACATCACAACATTTACTTAACGACAAGTCACCGAC 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 AGACTGTACCCATCCAGCCGCTGCACATCCAGCCGCTGCACATCACAACATTTACTTAACGACAAGTCACCGAC 1550

Qy   841 AACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTGCACGACGACGCC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 AACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTGCACGACGACGCC 1610

Qy   901 TGCCAGGGGCGATTCGGGAGGGGTGCCTCCCCGGTGTGCCTGAACGATGGCCGCATGACTTTGGTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 TGCCAGGGGCGATTCGGGAGGGGTGCCTCCCCGGTGTGCCTGAACGATGGCCGCATGACTTTGGTG 1670

Qy   961 GGCATCATCAGCTGGGGCTGGGCCTGGGCTGTGGACAGAGGATGTCCCGGGTGTACACAAAG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 GGCATCATCAGCTGGGGCTGGGCCTGGGCTGTGGACAGAGGATGTCCCGGGTGTACACAAAG 1730

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA 1065
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA 1775

RESULT 15
AAN90542
ID   AAN90542 standard; cDNA; 2560 BP.
XX
AC   AAN90542;
XX
DT   25-MAR-2003 (updated)
DT   05-JUN-1990 (first entry)
XX
DE   Plasmid pKG12 contg. DNA encoding human melanoma t-PA.
XX
KW   Human tissue plasminogen-activator gene; pKG12; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   sig_peptide 102..206
FT   /*tag= a
FT   mat_peptide ..1787
FT   /*tag= b
FT   /product=t-PA
XX
PN   EP297066-A.
XX
PD   28-DEC-1988.
XX
PF   14-JUN-1988; 88EP-0850207.
XX
PR   18-JUN-1987; 87SE-0002562.
XX
PA   (KABI) KABIGEN AB.
XX
PI   Pohl G, Hansson L, Loewenadler B;
XX
DR   WPI; 1989-001503/01.
DR   P-PSDB; AAP93716.
XX
PT   Modified tissue plasminogen-activator - having domains deleted and amino
PT   acid changes to increase biological half-life and reduce inactivation.
XX
PS   Disclosure; Fig 1; 18pp; English.
XX
CC   The sequence was obtained from cDNA prepd. from mRNA extracted from
CC   Bowes melanoma cells. The cDNA was used to construct a gene library which
CC   was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated
CC   and shown to comprise the whole coding region for human t-PA as well as
CC   102bp 5'flanking, 760 bp 3' flanking DNA and a poly A tail.
CC   See also AAN91608.
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;

Query Match   100.0%; Score 1065; DB 10; Length 2560;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC 785

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120

Wed Aug 20 09:29:43 2003

us-09-987-455-7.rng

```
Db   786 CTCACCGAGTCGGTGCCTCCGTCCCTGCCTCCTGCCTGAATTCCATGATGCCTGATAGGCAAGGTT 845
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 905

Qy   181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 965

Qy   241 GAGTACTGTGATGTGCCCTCCTGTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 GAGTACTGTGATGTGCCCTCCTGTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 1025

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCCAGGCTGCCATC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1026 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCCAGGCTGCCATC 1085

Qy   361 TTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCATCAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1086 TTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCATCAGC 1145

Qy   421 TCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCACCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1146 TCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCACCTG 1205

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1206 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 1265

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1266 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1325

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGTCCCGCTGTGCCCAGGAGAGCAGCGCTGGTCCGCACT 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1326 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGTCCCGCTGTGCCCAGGAGAGCAGCGCTGGTCCGCACT 1385

Qy   661 GTGTGCCCTTCCCCGGGGGACCTGCAGCTGCCGGACTCGCCGAGTGTGAGCTCTCCGGC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1386 GTGTGCCCTTCCCCGGGGGACCTGCAGCTGCCGGACTCGCCGAGTGTGAGCTCTCCGGC 1445

Qy   721 TACGGCAAGCATGAGGGCCTTGTGTCCTTCTATTCGGAGCGCGGCTGAAGGAGGCTCATGTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1446 TACGGCAAGCATGAGGGCCTTGTGTCCTTCTATTCGGAGCGCGGCTGAAGGAGGCTCATGTC 1505

Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTTAACAGAACAGTCACCGAC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1506 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTTAACAGAACAGTCACCGAC 1565

Qy   841 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACTTGCACGACGCC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1566 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACTTGCACGACGCC 1625

Qy   901 TGCCAGGGGCGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGCCCATGACTTTGGTG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1626 TGCCAGGGGCGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGCCCATGACTTTGGTG 1685

Qy   961 GGCATCATCAGCTGTGGGGCCTGGGCGTGTTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1686 GGCATCATCAGCTGTGGGGCCTGGGCGTGTTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1745

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTTGACAACATGCGACCGTGA 1065
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1746 GTTACCAACTACCTAGACTGGATTCGTTGACAACATGCGACCGTGA 1790
```

Search completed: August 18, 2003, 23:28:54
Job time : 376.192 secs

```
                GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    August 18, 2003, 22:51:23 ; Search time 83.3499 Seconds
                                      (without alignments)
                            5639.757 Million cell updates/sec

Title:          US-09-987-455-7
Perfect score:  1065
Sequence:       1 tctgagggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       569978 seqs, 22069156 residues

Total number of hits satisfying chosen parameters:   1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
         1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
         2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
         3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
         4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
         5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
         6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	Length	ID	Description
1	1065	100.0	1068	2	1068	US-08-811-949-44	Sequence 44, Appl
2	1065	100.0	1419	2	1419	US-08-811-949-62	Sequence 62, Appl
3	1065	100.0	1848	3	1848	US-08-814-412-10	Sequence 10, Appl
4	1065	100.0	2101	2	2101	US-08-811-949-42	Sequence 42, Appl
5	1065	100.0	2162	1	2162	US-08-119-512-3	Sequence 3, Appl
6	1065	100.0	2162	1	2162	US-08-488-015B-3	Sequence 3, Appl
7	1065	100.0	2162	1	2162	US-08-488-015B-25	Sequence 25, Appl
8	1063.4	99.8	1170	2	1170	US-08-811-949-64	Sequence 64, Appl
9	1063.4	99.8	1314	2	1314	US-08-811-949-48	Sequence 48, Appl
10	1063.4	99.8	1738	6	1738	5200340-1	Patent No. 5200340
11	1063.4	99.8	1955	2	1955	US-08-883-795A-39	Sequence 39, Appl
12	1063.4	99.8	2457	6	2457	5344773-1	Patent No. 5344773
13	1063.4	99.8	7360	5	7360	US-08-286-740-1	Sequence 1, Appl
14	1063.4	99.8	7360	5	7360	PCT-US95-09576-1	Sequence 1, Appl
15	1061.8	99.7	1068	2	1068	US-08-811-949-46	Sequence 46, Appl
16	1061.8	99.7	1314	2	1314	US-08-811-949-50	Sequence 50, Appl
17	1061.8	99.7	1974	2	1974	US-08-811-949-38	Sequence 38, Appl
18	1057	99.2	1068	2	1068	US-08-811-949-52	Sequence 52, Appl
19	1057	99.2	1314	2	1314	US-08-811-949-54	Sequence 54, Appl
20	1056.4	99.2	1068	1	1068	US-08-137-116-2	Sequence 2, Appl
21	1056.2	99.2	1137	4	1137	US-09-553-498-9	Sequence 9, Appl
22	1056.2	99.2	1137	4	1137	US-09-618-869-9	Sequence 9, Appl
23	1055.4	99.1	1170	2	1170	US-08-811-949-66	Sequence 66, Appl
24	1054.8	99.0	1065	1	1065	US-08-427-640-1	Sequence 1, Appli
25	1053.8	98.9	1314	4	1314	US-08-811-949-56	Sequence 56, Appl
26	1050	98.6	1068	6	1068	5223256-3	Patent No. 5223256
27	1046.8	98.3	1065	1	1065	US-08-427-640-5	Sequence 5, Appli
28	1041	97.7	1068	2		US-08-811-949-58	Sequence 58, Appl
29	1035.4	97.2	1068	1		US-08-427-640-3	Sequence 3, Appli
30	1033	97.0	1065	2		US-08-811-949-60	Sequence 60, Appl
31	953.4	89.5	2497	6		5185259-2	Patent No. 5185259
32	924	86.8	1068	2		US-08-427-640-7	Sequence 7, Appli
33	801	75.2	1163	2		US-08-558-269-5	Sequence 5, Appli
34	801	75.2	1163	3		US-09-410-882-5	Sequence 5, Appli
35	626.6	58.8	1727	6		5244676-4	Patent No. 5244676
36	625	58.7	1605	6		5244676-1	Patent No. 5244676
37	472	44.3	472	2		US-08-811-949-40	Sequence 40, Appl
38	417	39.2	453	6		5200340-3	Patent No. 5200340
39	264.4	24.8	1724	6		5200340-5	Patent No. 5200340
40	170.8	16.0	1236	1		US-07-957-039A-7	Sequence 7, Appli
41	170.8	16.0	1475	4		US-09-643-597-122	Sequence 122, App
42	170.8	16.0	1475	4		US-09-480-884A-122	Sequence 122, App
43	170.8	16.0	1475	4		US-09-542-615A-122	Sequence 122, App
44	170.8	16.0	1475	4		US-09-606-421B-122	Sequence 122, App
45	170.8	16.0	2301	6		5188829-2	Patent No. 5188829

ALIGNMENTS

```
RESULT 1
US-08-811-949-44

; Sequence 44, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;     ADDRESSEE: P.C.
;     STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;     CITY: ARLINGTON
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/811,949
;     FILING DATE: 05-MAR-1997
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: OBLON, NORMAN F.
;     REGISTRATION NUMBER: 24,618
;     REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-413-3000
;     TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 44:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1068 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..1065
US-08-811-949-44

  Query Match             100.0%;  Score 1065;  DB 2;  Length 1068;
```

```
Db    1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCACCGTGA 1068
```

```
RESULT 2
US-08-811-949-62
; Sequence 62, Application US/0881949
; Patent No. 5840533
; GENERAL INFORMATION:
;  APPLICANT: NIWA, MINEO
;  APPLICANT: SAITO, YOSHIMASA
;  APPLICANT: SASAKI, HITOSHI
;  APPLICANT: HAYASHI, MASAKO
;  APPLICANT: NOTANI, JOUJI
;  APPLICANT: KOBAYASHI, MASAKAZU
;  TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;  NUMBER OF SEQUENCES: 67
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 62:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 1419 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;  MOLECULE TYPE: DNA (genomic)
;  FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1416
US-08-811-949-62

Query Match       100.0%;  Score 1065;  DB 2;  Length 1419;
Best Local Similarity   100.0%;  Pred. No. 1.5e-291;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 414

Qy    61 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGCACTGCCCAGGCACTGGGCTGAATTCCATGATCCTGATAGGCAAGGTT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGCACTGCCCAGGCACTGGGCTGAATTCCATGATCCTGATAGGCAAGGTT 474

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCTGCAAACATAATTACTGCCGG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCTGCAAACATAATTACTGCCGG 534

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 594

Qy   241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGGCTGAGACAGTAGCAGCCCTCAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/814,412
  FILING DATE: 11-MAR-1997
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Jarrell Ph.D., Brenda H.
  REGISTRATION NUMBER: 39,223
  REFERENCE/DOCKET NUMBER: 0079571-0040
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617 248 5000
  TELEFAX: 617 248 4000
INFORMATION FOR SEQ ID NO: 10:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1848 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  IMMEDIATE SOURCE:
  CLONE: amplified t-PA clone
US-08-814-412-10

Query Match   100.0%; Score 1065; DB 3; Length 1848;
Best Local Similarity 100.0%; Pred. No. 1.7e-291;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    696 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGCACGCACAGC 755

Qy     61 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    756 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 815

Qy    121 TACACAGCACAGAACCCCAGTGCCAGCCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    816 TACACAGCACAGAACCCCAGTGCCAGCCACTGGGCCTGGGCAAACATAATTACTGCCGG 875

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGTGAAGAACCGCAGGCTGACGTGG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    876 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGTGAAGAACCGCAGGCTGACGTGG 935

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGGCTCGGCCTGAGACAGTACAGCCAGCCTCAG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    936 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGGCTCGGCCTGAGACAGTACAGCCAGCCTCAG 995

Qy    301 TTTCGCATCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    996 TTTCGCATCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 1055

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCCGGAGGTCGCCCGGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1056 TTTGCCAAGCACAGGAGGTCGCCCCGGAGGTCGCCCGGGGGCATACTCATCAGC 1115

Qy    421 TCCTGCTCGGATTCTTCTGCGCCCACTGCTTCCAGGAGGAGTTCCGCCCCACCACTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1116 TCCTGCTCGGATTCTTCTGCGCCCACTGCTTCCAGGAGGAGTTCCGCCCCACCACTG 1175

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1176 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA 1235

Qy    541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1236 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACAATGACATTGCG 1295

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTCCCGCGTGTGCCCAGGAGCAGCGTGTCCGCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1296 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTCCCGCGTGTGCCCAGGAGCAGCGTGTCCGCACT 1355
```

```
Db    595 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCCGACATGCAGCTGCTCAGAGACAGTACAGCAGCCGCCTCAG 654
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTCGCATCATCAAAGGAGGGCTCTTCGCCGACATGCAGCTGCTCCACCTTCCACCCTGCCATC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655 TTTCGCATCATCAAAGGAGGGCTCTTCGCCGACATGCAGCTGCTCCACCTTCCACCCTGCCATC 714

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGAGAGCGGCGTTCCTGTGCGGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715 TTTGCCAAGCACAGGAGGTCGCCCGAGAGCGGCGTTCCTGTGCGGGGGCATACTCATCAGC 774

Qy    421 TCCTGCTCTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 TCCTGCTCTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 834

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    835 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 894

Qy    541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 954

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCCGCGCCCGGAGCAGCGTGTGCCGCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955 CTGCTGCAGCTGAAATCGGATTCCTCCCCGCGCCCGGAGCAGCGTGTGCCGCACT 1014

Qy    661 GTGTGCGCCTTCCCCCGCCCCGGCCGACTGCCAGCTGCCCGACTCGACGGAGTGTGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 GTGTGCGCCTTCCCCCGCCCCGGCCGACTGCCAGCTGCCCGACTCGACGGAGTGTGAGCTCTCCGGC 1074

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGAGGCTCATGTC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGAGGCTCATGTC 1134

Qy    781 AGACTGTACCATCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 AGACTGTACCATCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1194

Qy    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1254

Qy    901 TGCCAGGGCGATTCGGAGGGCCCCCTGTGTCTGACGATGGCCGCCATGGCTTTGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 TGCCAGGGCGATTCGGAGGGCCCCCTGTGTCTGACGATGGCCGCCATGGCTTTGGTG 1314

Qy    961 GGCATCATCAGCTGGGGCCTGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 GGCATCATCAGCTGGGGCCTGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACAAAG 1374

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1419
```

RESULT 3
US-08-814-412-10
  Sequence 10, Application US/08814412
  ; Patent No. 6150141
GENERAL INFORMATION:
  APPLICANT: Jarrell Ph.D., Kevin A.
  TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
  NUMBER OF SEQUENCES: 46
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Choate, Hall & Stewart
  STREET: 53 State Street
  CITY: Boston
  STATE: MA
  COUNTRY: USA
  ZIP: 02109
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk

```
;  FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  25..1710
US-08-811-949-42

Query Match             100.0%;  Score 1065;  DB 2;  Length 2101;
Best Local Similarity   100.0%;  Pred. No. 1.8e-291;
Matches 1065;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCACCGTGGCACGCACCACAGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   649 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCACCGTGGCACGCACCACAGC 708

Qy    61 CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   709 CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 768

Qy   121 TACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   769 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 828

Qy   181 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   829 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG 888

Qy   241 GAGTACTGTGATGTGCCCTGCTCACCTGCGCGCCTGAGACAGTACAGCCAGCCTCAG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   889 GAGTACTGTGATGTGCCCTGCTCACCTGCGCGCCTGAGACAGTACAGCCAGCCTCAG 948

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCGCCGGAGAGCGGTTCCTGTGCGCGGGGCATCGCCATC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   949 TTTCGCATCAAAGGAGGGCTCTTCGCGCCGGAGAGCGGTTCCTGTGCGCGGGGCATCGCCATC 1008

Qy   361 TTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTCGCCGGAGAGCGGGGCATACTACTCATCAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1009 TTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTCGCCGGAGAGCGGGGCATACTACTCATCAGC 1068

Qy   421 TCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1069 TCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG 1128

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGTGTCCCTGGCTGGTGTCCCTGGCCAGAAATTTGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1129 ACGGTGATCTTGGGCAGAACATACCGGTGTCCCTGGCTGGTGTCCCTGGCCAGAAATTTGAA 1188

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTACGACAATGACATTGCG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1189 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTACGACAATGACATTGCG 1248

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTTCCGCGGATTCGGATTCGTGTCCGGAGGAGCAGCGGTGTCCGCACT 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1249 CTGCTGCAGCTGAAATCGGATTCGTTCCGCGGATTCGGATTCGTGTCCGGAGGAGCAGCGGTGTCCGCACT 1308

Qy   661 GTGTGCCTTCCCCGGCGGACTGGAGTGCGCCGGAGTGCGCCGGAGTGTGAGCTCTCCGGC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1309 GTGTGCCTTCCCCGGCGGACTGGAGTGCGCCGGAGTGCGCCGGAGTGTGAGCTCTCCGGC 1368

Qy   721 GTGTGCCTTCCCCGCCGCAGTGGCCTCCAGCTGCAGTGGCAGTGGCCTCCAGCTGGTCTCCGGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1369 GTGTGCCTTCCCCGCCGCAGTGGCCTCCAGCTGCAGTGGCAGTGGCCTCCAGCTGGTCTCCGGC 1428

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1429 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1488

Qy   841 AACATGCTGTGTGCTGCTGAGACACTCGGAGCGGCGGGCCCAAACTTGCACGACGCC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1489 AACATGCTGTGTGCTGCTGAGACACTCGGAGCGGCGGGCCCAAACTTGCACGACGCC 1548

Qy   901 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCCATGAGACTTTGGTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1549 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCCATGAGACTTTGGTG 1608
```

```
Qy   661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACGGAGTGTGAGCTCTCCGGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1356 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACGGAGTGTGAGCTCTCCGGC 1415

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTCATTCGAGCGGCTGAGGAGGCTCATGTC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1416 TACGGCAAGCATGAGGCCTTGTCTCCTTCATTCGGGAGCGGCTGAGGAGGCTCATGTC 1475

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1476 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1535

Qy   841 AACATGCTGTGTGCTGCTGAGACACTCGGAGCGGCGGGCCCAAACTTGCACGACGCC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1536 AACATGCTGTGTGCTGCTGAGACACTCGGAGCGGCGGGCCCAAACTTGCACGACGCC 1595

Qy   901 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCCATGAGACTTTGGTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1596 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCCATGAGACTTTGGTG 1655

Qy   961 GGCATCATCAGCTGGGGCTGGGCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTGTACACAAAG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1656 GGCATCATCAGCTGGGGCTGGGCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTGTACACAAAG 1715

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
           ||||||||||||||||||||||||||||||||||||||||
Db  1716 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1760
```

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCGCTACCGTGCACGCACAGC    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCGCTACCGTGCACGCACAGC   765

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCAGGCACTGGGCTTCCATGATCCTGATAGGCAAGGTT   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCAGGCACTGGGCTTCCATGATCCTGATAGGCAAGGTT   825

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG   885

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGAGCTGACGTGG   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGAGCTGACGTGG   945

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGGCTGAGACAGTACAGCCAGCCTCAG   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGGCTGAGACAGTACAGCCAGCCTCAG  1005

Qy    301 TTTCGCCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TTTCGCCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC  1065

Qy    361 TTTGTCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCGTGTGCGGGGCATACTCATCATCAGC   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCGTGTGCGGGGCATACTCATCATCAGC  1125

Qy    421 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCGTGCCCCCCACCACCTG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCGTGCCCCCCACCACCTG  1185

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA  1245

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG  1305

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGAGCAGCGTGGTCCGCACT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGAGCAGCGTGGTCCGCACT  1365

Qy    661 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACTGGACCTGGACCTGGAGCTCTCCGGC   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACTGGACCTGGACCTGGAGCTCTCCGGC  1425

Qy    721 TACGGCAAGCAATGGAGGGCTTGCTGTCCCTTCTATTCGGCGGCTGAACGGAGGAGGCTCATGTC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 TACGGCAAGCAATGGAGGGCTTGCTGTCCCTTCTATTCGGCGGCTGAACGGAGGAGGCTCATGTC  1485

Qy    781 AGACTGTACCCATCCAGCGCGTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 AGACTGTACCCATCCAGCGCGTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  1545

Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACTTGCACGACGCC   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACTTGCACGACGCCC  1605

Qy    901 TGCCAGGCGGATTCGGAGGACCCCTGTGTCTGAACGATGGCCCCATGACGTCTGGTG   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 TGCCAGGCGGATTCGGAGGACCCCTGTGTCTGAACGATGGCCCCATGACGTTTGGTG  1665

Qy    961 GGCATCATCAGCTGGGGCCTGGGCTGCTGTGGACAAGAGATGTCCCGGGTGTGTACACAAAG  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 GGCATCATCAGCTGGGGCCTGGGCTGCTGTGGACAAGAGATGTCCCGGGTGTGTACACAAAG  1725

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1770
```

RESULT 5
US-08-119-512-3
; Sequence 3, Application US/08119512
; Patent No. 5498531
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,512
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUI-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 82..334
; OTHER INFORMATION: /product= "Signal Sequence and
; OTHER INFORMATION: Finger-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 335..447
; OTHER INFORMATION: /product= "EGF-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448..714
; OTHER INFORMATION: /product= "Kringle-1 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 715..972
; OTHER INFORMATION: /product= "Kringle-2 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 973..2162
; OTHER INFORMATION: /product= "Catalytic domain"
US-08-119-512-3

Query Match   100.0%; Score 1065; DB 1; Length 2162;
Best Local Similarity   100.0%; Pred. No. 1.8e-291;

```
Db   766 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 TACACAGCACAGAACCCAGTGCCTGGCACTGGGCCTGGCACTGGGCCAAACATAATTACTGCCGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   826 TACACAGCACAGAACCCAGTGCCTGGCACTGGGCCTGGCACTGGGCCAAACATAATTACTGCCGG 885

Qy   181 AATCCTGATGGGGATGCCAAGCCTGTGCCACGTGCTGAAGAACCGGAGGCTGACGTGG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   886 AATCCTGATGGGGATGCCAAGCCTGTGCCACGTGCTGAAGAACCGGAGGCTGACGTGG 945

Qy   241 GAGTACTGTGATGTGCCTCCTGCTCACCTGGGGCTGAGACAGTACAGCCAGCCTCAG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   946 GAGTACTGTGATGTGCCTCCTGCTCACCTGGGGCTGAGACAGTACAGCCAGCCTCAG 1005

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1006 TTTCGCATCAAAGGAGGGCTCTTCGCCACATCGCCTCCACCCCTGGCAGGCTGCCATC 1065

Qy   361 TTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1066 TTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 1125

Qy   421 TCCTGCTGGATTCTCTCGCCCCACTGCTTCCGGAGGAGGTTTCCGCCCCACCACCTG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1126 TCCTGCTGGATTCTCTCGCCCCACTGCTTCCGGAGGAGGTTTCCGCCCCACCACCTG 1185

Qy   481 ACGGTGATCTTGGCAGAACATACCGGGTGGTTCCTGGCGAGGAGGAGCAGAAATTTGAA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1186 ACGGTGATCTTGGCAGAACATACCGGGTGGTTCCTGGCGAGGAGGAGCAGAAATTTGAA 1245

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1305

Qy   601 CTGCTGCAGCTGAAATGGATTCGTCCGCGTGCTGCCAGGAGCAGCAGGCGTGGTCCGCACT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1306 CTGCTGCAGCTGAAATGGATTCGTCCGCGTGCTGCCAGGAGCAGCAGGCGTGGTCCGCACT 1365

Qy   661 GTGTGCCTTCCCCGCGGACTGCAGCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1366 GTGTGCCTTCCCCGCGGACTGCAGCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 1425

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1426 TACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1485

Qy   781 AGACTGTACCCATCCAGCGACACAGTCGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1486 AGACTGTACCCATCCAGCGACACAGTCGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1545

Qy   841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGCGCAAACTTGCACGACGCC 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1546 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGCGCAAACTTGCACGACGCC 1605

Qy   901 TGCCAGGGCGATTCGGGAGGCCCCCTGGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1606 TGCCAGGGCGATTCGGGAGGCCCCCTGGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1665

Qy   961 GGCATCATCAGCTGGGGCCTGGGCTGTGGACGAAGGATGTCCCGGGTGTGTACAAAG 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1666 GGCATCATCAGCTGGGGCCTGGGCTGTGGACGAAGGATGTCCCGGGTGTGTACAAAG 1725

Qy  1021 GTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCGTGA 1065
             ||||||||||||||||||||||||||||||||||||||||||||
Db  1726 GTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCGTGA 1770
```

RESULT 7
US-08-488-015B-25
; Sequence 25, Application US/08488015B
; Patent No. 5780272
; GENERAL INFORMATION:

RESULT 6
US-08-488-015B-3
; Sequence 3, Application US/08488015B
; Patent No. 5780272
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,015B
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUV-008.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 82..334
; OTHER INFORMATION: /product= "Signal Sequence and
Patent No. 5780272
; OTHER INFORMATION: Finger-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 335..447
; OTHER INFORMATION: /product= "EGF-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448..714
; OTHER INFORMATION: /product= "Kringle-1 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 715..972
; OTHER INFORMATION: /product= "Kringle-2 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 973..2162
; OTHER INFORMATION: /product= "Catalytic domain"
US-08-488-015B-3

Query Match   100.0%; Score 1065; DB 1; Length 2162;
Best Local Similarity 100.0%; Pred. No. 1.8e-291;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TCTGGAGGAAAACAGTGACTGCTACTTTGGGAATGGGCTCAGCCTTACCGTGGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  706 TCTGGAGGAAAACAGTGACTGCTACTTTGGGAATGGGCTCAGCCTTACCGTGGCACGCACAGC 765

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
```

```
Db  1186 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGCAGAAATTTGAA 1245
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 GTCGAAAAATACATTGTCCATAAGGAAATCGATGATGACATTCGATGATGACACATTGCG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1246 GTCGAAAAATACATTGTCCATAAGGAAATCGATGATGACACTTCGATGATGACACATTGCG 1305
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGGAGCAGCGTGGTCCGCACT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1306 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGGAGCAGCGTGGTCCGCACT 1365
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 GTGTGGCCTTCCCCCGGCCAGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1366 GTGTGGCCTTCCCCCGGCCAGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 1425
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 TACGGCAAGCATGAGGCCCTGTCTCCTTTCGGAGCGGCTGAAGGAGGCTCATGTC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1426 TACGGCAAGCATGAGGCCCTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1485
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGAACAGTCACCGAC 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1486 AGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGAACAGTCACCGAC 1545
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 AACATGCTGTGTGCTGGAGACACTCGGACGCGGCGGCCCCCAGGCAAACTTGCACGACGCC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1546 AACATGCTGTGTGCTGGAGACACTCGGACGCGGCGGCCCCCAGGCAAACTTGCACGACGCC 1605
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 TGCCAGGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1606 TGCCAGGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1665
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961 GGCATCATCAGCTGGGGCTGGGGCTGTGAGCAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1666 GGCATCATCAGCTGGGGCTGGGGCTGTGAGCAGAAGGATGTCCCGGGTGTGTACACAAAG 1725
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
              |||||||||||||||||||||||||||||||||||||||||||||
Db  1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1770
```

RESULT 8
US-08-811-949-64
; Sequence 64, Application US/0811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0

; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,015B
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUV-008.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2162 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
; US-08-488-015B-25

Query Match   100.0%; Score 1065; DB 1; Length 2162;
Best Local Similarity   100.0%; Pred. No. 1.8e-291;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   706 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    61 CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   766 CTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 TACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCCAAACATAATTAATTACTGCCGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   826 TACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCCAAACATAATTAATTACTGCCGG 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   886 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 GAGTACTGTGATGTGCCCTTCCACCTGCGGCCTGAGACGAGTACAGCCAGCCTCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   946 GAGTACTGTGATGTGCCCTTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTTCGCCACCCCTGCCAGGCTGCCATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1006 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTTCGCCACCCCTGCCAGGCTGCCATC 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1066 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 TCCTGCTGGATTCTCTGCGCCCACTGCTTCGGCGGGAGGTTTCCGCCCCACCACCTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1126 TCCTGCTGGATTCTCTGCGCCCACTGCTTCGGCGGGAGGTTTCCGCCCCACCACCTG 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGCAGAAATTTGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  703-413-3000
;        TELEFAX:    703-413-2220
; INFORMATION FOR SEQ ID NO:    64:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:   1170 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  double
;        TOPOLOGY:  circular
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;        NAME/KEY:  CDS
;        LOCATION:  1..1167
US-08-811-949-64
```

Query Match   99.8%;   Score 1063.4; DB 2; Length 1170;
Best Local Similarity   99.9%;   Pred. No. 4e-291;
Matches 1064; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 165

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTTCCATGATCCTGATAGGCAAGGTT 225

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 285

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 345

Qy    241 GAGTACTGTGATGTGCCCTCGCTCCACCTGCGGGCCTGCCAGTACGACCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 GAGTACTGTGATGTGCCCTCGCTCCACCTGCGGGCCTGCCAGTACGACCAGCCTCAG 405

Qy    301 TTTCGCATCAAAGAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGCGAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 TTTCGCATCAAAGAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGCGAGGCTGCCATC 465

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGTCGGGGGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGTCGGGGGCATACTCATCAGC 525

Qy    421 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCGGAGGAGAGGTTTCCGCCCCCACCACCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCGGAGGAGAGGTTTCCGCCCCCACCACCTG 585

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA 645

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 705

Qy    601 CTGCTGCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 CTGCTGCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 765

Qy    661 GTGTGCCTTCCCCCGGCGGACCTGCCAGCTGCCGCCGGAGTGTGAGCTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 GTGTGCCTTCCCCCGGCGGACCTGCCAGCTGCCGCCGGAGTGTGAGCTCTCCGGC 825

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 885

Qy    781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 945

Qy    841 AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCGGGGCCCCAGCCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCGGGGCCCCAGCCAAACTTGCACGACGCC 1005

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGCCCGATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGCCGATGACTTTGGTG 1065

Qy    961 GGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 GGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1125

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1170
```

RESULT 9
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 48:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1314 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1311
US-08-811-949-48

Query Match   99.8%;   Score 1063.4; DB 2; Length 1314;
Best Local Similarity   99.9%;   Pred. No. 4.2e-291;
Matches 1064; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
```

5200340-1
;Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
;SEQ ID NO:1:
; LENGTH: 1738
5200340-1

Query Match   99.8%;   Score 1063.4;   DB 6;   Length 1738;
Best Local Similarity 99.9%;   Pred. No. 4.8e-291;
Matches 1064; Conservative   0; Mismatches   0; Indels 1; Gaps   0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACCAGC    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    640 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACCAGC   699

Qy     61 CTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    700 CTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT   759

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGCCTGGCAAACATAATTACTGCCGG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    760 TACACAGCACAGAACCCCAGTGCCAGGCACTGGCCTGGCAAACATAATTACTGCCGG   819

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    820 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG   879

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    880 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   939

Qy    301 TTTCGCATCAAAGAGAGGGCTCTTCGCCGACATGCCGGTGTTCCGCCCCACCACCTG   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    940 TTTCGCATCAAAGAGAGGGCTCTTCGCCGACATGCCGGTGTTCCGCCCCACCACCTG   999

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCTGTCGCGGAGCGGTCTGCCATC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1000 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCTGTCGCGGAGCGGTCTGCCATC  1059

Qy    421 TCCTGCTGGATTCTCTGCCCCACTGCTTCGCCAGGAGAGGTTTCCGCCCCACCACCTG   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 TCCTGCTGGATTCTCTGCCCCACTGCTTCGCCAGGAGAGGTTTCCGCCCCACCACCTG  1119

Qy    481 ACGGTGATCTTGGGCAGAACATACGGGTGGTCCTGGGAGGAGGAGCAGAAATTTGAA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1120 ACGGTGATCTTGGGCAGAACATACGGGTGGTCCTGGGAGGAGGAGCAGAAATTTGAA  1179

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATTGCCG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1180 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATTGCCG  1239

Qy    601 CTGCTGCAGCTGAAATCGGATTCGGTGTCCCGGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1240 CTGCTGCAGCTGAAATCGGATTCGGTGTCCCGGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT  1299

Qy    661 GTGTGCCTTCCCCGCGACCTGCAGCTGCGACCTGGAGGTGTGAGCTCCCGGC   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1300 GTGTGCCTTCCCCGCGACCTGCAGCTGCGACCTGGAGGTGTGAGCTCCCGGC  1359

Qy    721 TACGGCAAGCATGAGGCCTTGCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1360 TACGGCAAGCATGAGGCCTTGCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  1419

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1420 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  1479
```

Qy   121 TACACAGCACAGAACCCAGTGCCAGCACTGGGCCTGGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835 TACACAGCACAGAACCCAGTGCCAGCACTGGGCCTGGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 894

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGGGCCACGTGCTGGGCCACGTGCTGGGCGACGTGG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGGGCCACGTGCTGGGCGACGTGG 954

Qy   241 GAGTACTGTGATGTGCCTCCTGTCTCCACCTGGGCCTGGGCCTGGGCCTCAGCCTCAG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955 GAGTACTGTGATGTGCCTCCTGTCTCCACCTGGGCCTGGGCCTGAGCACCAGCCTCAG 1014

Qy   301 TTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1015 TTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 1074

Qy   361 TTTGCCAAGCACAGCAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCCATACTCATCAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1075 TTTGCCAAGCACAGCAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCCATACTCATCAGC 1134

Qy   421 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCGGAGAGAGGTTTCCGCCCACCACCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1135 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCGGAGAGAGGTTTCCGCCCACCACCTG 1194

Qy   481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1195 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 1254

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAATGACACTTACGACAATGACATTGCG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1255 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAATGACACTTACGACAATGACATTGCG 1314

Qy   601 CTGCTGCAGCTGAAATCGATTCGTCCCGTGCCCGGACTGGACCGGAGTGGAGCTCTCCCGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1315 CTGCTGCAGCTGAAATCGATTCGTCCCGTGCCCGGACTGGACCGGAGTGGAGCTCTCCCGC 1374

Qy   661 GTGTGCCTTCCCCGGGCCAGCTGCAGCTGGACCTGGACCGGAGTGGAGCTGTGAGCTCCGC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1375 GTGTGCCTTCCCCGGGCCAGCTGCAGCTGGACCTGGACCGGAGTGGAGCTGTGAGCTCCGC 1434

Qy   721 TACGGCAAGCAATGAGGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1435 TACGGCAAGCAATGAGGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 1494

Qy   781 AGACTGTACCCATCCAGCGCCGCTGCACATCACAACATTTACTTAACAGACAGTCACCGAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1495 AGACTGTACCCATCCAGCGCCGCTGCACATCACAACATTTACTTAACAGACAGTCACCGAC 1554

Qy   841 AACATGTCTGTGCTGAGACACTCGGAGCGGCGGCCCAGGCAAACTTGCACGACGCC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1555 AACATGTCTGTGCTGAGACACTCGGAGCGGCGGCCCAGGCAAACTTGCACGACGCGC 1614

Qy   901 TGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTCTGAACGATGCCGCATGACTTGGTG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1615 TGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTCTGAACGATGCCGCATGACTTTGGTG 1674

Qy   961 GGCATCATCAGCTCGGGTGCCTCGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1675 GGCATCATCAGCTCGGGTGCCTCGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1734

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
         |||||||||||||||||||||||||||||||||||||||||
Db 1735 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1779

Qy   841 AACATGTGTGTGCTGAGACACTCGAGCGGCGGGGGCCCAGGCAAACTTGCACGACGCC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1480 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGGCCCGGAGCGGCGGGGGCCCCAGCAAACTTGCACGACGCGC 1539

Qy   901 TGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTCTGAACGATGCCGCATGACTTGGTG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1540 TGCCAGGGGCGATTCGGGAGGCGCCCCCTGTGTCTGAACGATGGCGCATGACTTTGGTG 1599

Qy   961 GGCATCATCAGCTCGGGTGCCTGTGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1600 GGCATCATCAGCTCGGGTGCCTGTGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1659

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
         |||||||||||||||||||||||||||||||||||||||||
Db 1660 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1704

RESULT 11
US-08-883-795A-39
; Sequence 39, Application US/08883795A
; Patent No. 5985607
GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1955 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
US-08-883-795A-39

Query Match   99.8%; Score 1063.4; DB 2; Length 1955;
Best Local Similarity 99.8%; Pred. No. 5e-291;
Matches 1064; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCAGCTGGCACGCACAGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 715 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCAGCTGGCACGCACAGC 774

Qy   61 CTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGTT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 775 CTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGTT 834

RESULT 12
534773-1
; Patent No. 534773
; APPLICANT: WEI, CHA-MER; HSIUNG, NANCY; REDDY, VERMURI B.;
; LEMONTT, JEFFREY F.; DACKOWSKI, WILLIAM; DOUGLAS, RICHARD;
; COLE, EDWARD S.; PURCELL JR., RICHARD D.; LAU, DAVID TAI-YUI
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN

```
;ACTIVATOR PRODUCED BY RECOMBINANT DNA
;   NUMBER OF SEQUENCES: 6
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/07/782,686
;     FILING DATE: 01-OCT-1985
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 656,770
;     FILING DATE: 01-OCT-1984
;SEQ ID NO:1:
;   LENGTH: 2457
5344773-1

Query Match         99.8%;   Score 1063.4;  DB 6;  Length 2457;
Best Local Similarity  99.9%;  Pred. No. 5.5e-291;
Matches 1064;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  638 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 697

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCCGTGGAATTCCTGATCCTGATAGGCAAGGTT 120
        ||||||||||||||||||||||||||||||||||||||||||||
Db  698 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCCGTGGAATTCCTGATCCTGATAGGCAAGGTT 757

Qy  121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  758 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 817

Qy  181 AATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  818 AATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 877

Qy  241 GAGTACTGTGATGCTGCCCCTCGTGCGGGCTACAGCACAGTACAGCCAGCCTCAG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  878 GAGTACTGTGATGCTGCCCCTCGTGCGGGCTACAGCACAGTACAGCCAGCCTCAG 937

Qy  301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACAATCGCCTCCCACCCCCTGGCAGGCTGCCATC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  938 TTTCGCATCAAAGGAGGGCTCTTCGCCGACAATCGCCTCCCACCCCCTGGCAGGCTGCCATC 997

Qy  361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  998 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 1057

Qy  421 TCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCGGCCCCCCACCACCTG 480
        ||||||||||||||||||||||||||||||||||||||||||||
Db 1058 TCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCGGCCCCCCACCACCTG 1117

Qy  481 ACGGTGATCTTGGCAGAACATACGGGTGTGTCCCTGGCGAGGAGCAGAAATTTGAA 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1118 ACGGTGATCTTGGCAGAACATACGGGTGTGTCCCTGGCGAGGAGCAGAAATTTGAA 1177

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1178 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1237

Qy  601 CTGCTGCAGCTGAAATCGGAATTCGTTCCCGCTGTGCCCAGGAGACAGCAGGTGGTCCGCACT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1238 CTGCTGCAGCTGAAATCGGAATTCGTTCCCGCTGTGCCCAGGAGACAGCAGGTGGTCCGCACT 1297

Qy  661 GTGTGCCTTCCCCCGGGCGGGGACTGCAGCTGACCGGACTGGGAGTGTGACGGTCTCCGGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1298 GTGTGCCTTCCCCCGGGCGGGGACTGCAGCTGACCGGACTGGGAGTGTGACGGTCTCCGGC 1357

Qy  721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGAGGCTCATGTC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1358 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGAGGCTCATGTC 1417

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACACAACATTTACTTAACGACAGTCACCGAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1418 AGACTGTACCCATCCAGCCGCTGCACATCACACAACATTTACTTAACGACAGTCACCGAC 1477

Qy  841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCCAGGCCAAACTTGCGACGACGCC 900
```

```
Db 1478 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCCAGGCCAAACTTGCACGACGCC 1537
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901 TGGCAGGGCGGATTCGGGAGGAGGCCCCTGTGTGTCTGACGCGGATGGCGCCGCATTGGTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1538 TGGCAGGGCGGATTCGGGAGGAGGCCCCTGTGTGTCTGACGCGGATGGCGCCGCATTGGTG 1597

Qy  961 GGCATCATCAGCTGGGGCCTGGGCTGTGCGGATGGACAGAAGAGGATGTCCCGGGTGTGTACACAAAG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1598 GGCATCATCAGCTGGGGCCTGGGCTGTGCGGATGGACAGAAGAGGATGTCCCGGGTGTGTACACCAAG 1657

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
        |||||||||||||||||||||||||||||||||||||||||||||
Db 1658 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1702
```

```
RESULT 13
US-08-286-740-1
; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
;   APPLICANT: Crowley, Craig W.
;   TITLE OF INVENTION:  METHOD FOR SELECTING HIGH-EXPRESSING
;   TITLE OF INVENTION:  HOST CELLS
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 460 Point San Bruno Blvd
;     CITY: South San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/286,740
;     FILING DATE: 05-AUG-1994
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Lee, Wendy M.
;     REGISTRATION NUMBER:
;     REFERENCE/DOCKET NUMBER: 798
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415/225-1994
;     TELEFAX: 415/952-9881
;     TELEX: 910/371-7168
;INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7360 bases
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-286-740-1

Query Match         99.8%;   Score 1063.4;  DB 1;  Length 7360;
Best Local Similarity  99.9%;  Pred. No. 8.7e-291;
Matches 1064;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2246 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCCACAGC 2305

Qy   61 CTCACCGAGTGCGGGTGCGTGCCTCCTGCCTCCTGCTAGCTTCCATGATCCTGATAGGCAAGGTT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2306 CTCACCGAGTGCGGGTGCGTGCCTCCTGCCTCCTGCTAGCTTCCATGATCCTGATAGGCAAGGTT 2365

Qy  121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTTGGGCAAACATAATTACTGCCGG 180
```

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09576
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286740
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 798PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9981
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7360 bases
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US95-09576-1

Query Match   99.8%;   Score 1063.4;   DB 5;   Length 7360;
Best Local Similarity 99.9%;   Pred. No. 8.7e-291;
Matches 1064;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 2305

Qy   61 CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 2365

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2366 TACACAGCACAGAACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG 2425

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGTGTGAGCTGTGAGCTGTGACGTGG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 AATCCTGATGGGGATGCCAAGCCCTGGTGTGTGAGCTGTGAGCTGTGACGTGG 2485

Qy   241 GAGTACTGTGATGTGCCCCTGCCTGAGACAGTAGCAGTCACCGAC 300
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2486 GAGTACTGTGATGTGCCCCTGCCTGAGACAGTAGCAGTCACCGAC 2545

Qy   301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGCTGCCATC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2546 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGCTGCCATC 2605

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCGCCCACCACTCATCAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCGCCCACCACTCATCAGC 2665

Qy   421 TCCTGCTGGATTCTCTCGCCCACTGCTTCCGGAGAGAGTTTCGCCCACCACCTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2666 TCCTGCTGGATTCTCTCGCCCACTGCTTCCGGAGAGAGTTTCGCCCACCACCTG 2725

Qy   481 ACGGTGATCTTGGCAGAACATACCGGGTGCCTGGCGACATTTGAA 540
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2726 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGACATTTGAA 2785

Db   2366 TACACAGCACAGAACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG 2425

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGGGCAAGCGGTGG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGGGCAAGCGGTGG 2485

Qy   241 GAGTACTGTGATGTGCCCCTGCCTCCACCTGGGCGCCTGAGACAGTAGCAGTCACCGAC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 GAGTACTGTGATGTGCCCCTGCCTCCACCTGGGCGCCTGAGACAGTAGCAGTCACCGAC 3085

Qy   841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCGCC 3145

Qy   901 TGCCAGGGCGATTCGGGAGGCCTTGGGCTGTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 TGCCAGGGCGATTCGGGAGGCCTTGGGCTGTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 3205

Qy   961 GGCATCATCAGCTGGGGCCTGGGCCTGGGCGTGTGACAGAAGAAGATGTCCGGGTGTGTACAAAAG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3206 GGCATCATCAGCTGGGGCCTGGGCCTGGGCGTGTGACAGAAGAAGATGTCCGGGTGTGTACAAAAG 3265

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 3310

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCACGCGTGGTCCGCACT 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCACGCGTGGTCCGCACT 2905

Qy   661 GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGACTGGACCGGAGTGTGAGCTCTCGGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGACTGGACCGGAGTGTGAGCTCTCGGC 2965

Qy   721 TACGGCAAGCATCAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 3025

Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 3085

RESULT 14
PCT-US95-09576-1
; Sequence 1, Application PC/TUS9509576
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```
                     TELEPHONE:  703-413-3000
                     TELEFAX:    703-413-2220
               INFORMATION FOR SEQ ID NO:  46:
               SEQUENCE CHARACTERISTICS:
                     LENGTH:  1068 base pairs
                     TYPE:  nucleic acid
                     STRANDEDNESS:  double
                     TOPOLOGY:  circular
               MOLECULE TYPE:  DNA (genomic)
               FEATURE:
                     NAME/KEY:  CDS
                     LOCATION:  1..1065
 US-08-811-949-46

 Query Match            99.7%;  Score 1061.8;  DB 2;  Length 1068;
 Best Local Similarity  99.8%;  Pred. No. 1.1e-290;
 Matches 1063;  Conservative  0;  Mismatches  0;  Indels  2;  Gaps  0;

Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 63

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 123

Qy  121 TACAGACAGAACCCCAGTGCCAGGGACTGGGCCTGGCAAACATAATTACTGCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124 TACAGACAGAACCCCAGTGCCAGGGACTGGGCCTGGCAAACATAATTACTGCCG 183

Qy  181 AATCCTGATGGGGATGGCAAGCCTGGTGCCACGTGTGAAGAACCGAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184 AATCCTGATGGGGATGGCAAGCCTGGTGCCACGTGTGAAGAACCGAGGCTGACGTGG 243

Qy  241 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGGCTGAGACAGTACAGCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGGCTGAGACAGTACAGCAGCCTCAG 303

Qy  301 TTTCGCATCAAGGAGGGCTCTTCGCACGACATGCGCCTCCACCCCTGGCAGGCTGCCATC 360
          |||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||||
Db  304 TTTCGCATCATCAGGAGGGCTCTTCGCACGACATCGCCTCCACCCCTGGCAGGCTGCCATC 363

Qy  361 TTTGCCAAGCACAGGAGGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364 TTTGCCAAGCACAGGAGGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 423

Qy  421 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 483

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCTGGAGGAGGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCTGGAGGAGGCAGAAATTTGAA 543

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 603

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTCGCCGGAGGAGCAGCGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604 CTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTCGCCGGAGGAGCAGCGTGGTCCGCACT 663

Qy  661 GTGTGCCTTCCCCGGCGGGGACCTGCAGCTGCGGACTGGACGGAGTGTGAGCGTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664 GTGTGCCTTCCCCGGCGGGGACCTGCAGCTGCGGACTGGACGGAGTGTGAGCGTCTCCGGC 723

Qy  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  724 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAAGGAGGCTCATGTC 783

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 843
```

```
Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2786 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 2845

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTCGCCGGAGGAGCAGCGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2846 CTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTCGCCGGAGGAGCAGCGTGGTCCGCACT 2905

Qy  661 GTGTGCCTTCCCCGGCGGGGACCTGCAGCTGCGGACTGGACGGAGTGTGAGCGTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2906 GTGTGCCTTCCCCGGCGGGGACCTGCAGCTGCGGACTGGACGGAGTGTGAGCGTCTCCGGC 2965

Qy  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2966 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAAGGAGGCTCATGTC 3025

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3026 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 3085

Qy  841 AACATGCTGTGTCTGAGACACTCGGAGACACTTGACGAGACGCC 900
          |||||||||||||||||||||||||||||||||||||||||||||
Db 3086 AACATGCTGTGTCTGAGACACTCGGAGACACTTGACGAGACGCC 3145

Qy  901 TGGCCAGGGCGAATTCGGGAGGGCCCCCTGTGTCTGACGATGGCCGCCATGACTTTCGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3146 TGGCCAGGGCGAATTCGGGAGGGCCCCCTGTGTCTGACGATGGCCGCCATGACTTTCGGTG 3205

Qy  961 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAGGATGTCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3206 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAGGATGTCCGGGTGTGTACACCAAG 3265

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db 3266 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 3310
```

RESULT 15
US-08-811-949-46:
; Sequence 46, Application US/08811949
; Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
    ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
    ADDRSSEE: P.C.
    STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
    CITY: ARLINGTON
    STATE: VA
    COUNTRY: USA
    ZIP: 22202
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/811,949
    FILING DATE: 05-MAR-1997
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: OBLON, NORMAN F.
    REGISTRATION NUMBER: 24,618
    REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:

```
Qy    841 AACATGCTGTGTGTCTGAGACACTCGGAGCGGCGGGCGCCCCAGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 AACATGCTGTGTGTCTGGAGACACTCGGAGCGGCGGGCGCCCCAGGCAAACTTGCACGACGCC 903

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 963

Qy    961 GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1023

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1068
```

Search completed: August 19, 2003, 08:35:12
Job time : 85.3499 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 300.529 Seconds
(without alignments)
7930.701 Million cell updates/sec

Title: US-09-987-455-7
Perfect score: 1065
Sequence: 1 tctgaggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
```
1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1065	100.0	1065	11	US-09-987-455-4	Sequence 4, Appli
2	1065	100.0	1065	11	US-09-987-455-7	Sequence 7, Appli
3	1063.4	99.8	1689	9	US-09-969-271-6	Sequence 6, Appl
4	1063.4	99.8	2509	14	US-10-193-656-7	Sequence 7, Appli
5	1063.4	99.8	2519	9	US-09-969-271-5	Sequence 5, Appli
6	1062	99.7	1128	11	US-09-987-455-2	Sequence 2, Appli
7	1062	99.7	1128	11	US-09-987-455-5	Sequence 5, Appli
8	1061.8	99.7	2641	10	US-09-974-298-144	Sequence 144, App
9	230.2	21.6	329	12	US-10-007-926A-433	Sequence 433, App
10	191.6	18.0	1212	10	US-09-880-503-15	Sequence 15, Appl
11	185	17.4	1415	14	US-10-198-846-12748	Sequence 12748, A
12	170.8	16.0	1475	9	US-09-735-705-122	Sequence 122, App
13	170.8	16.0	1475	10	US-09-850-716A-122	Sequence 122, App
14	170.8	16.0	1475	10	US-09-897-778-122	Sequence 122, App
15	170.8	16.0	1475	11	US-09-466-396A-122	Sequence 122, App
16	170.8	16.0	1475	12	US-10-117-982-122	Sequence 122, App
17	170.8	16.0	1475	12	US-10-101-510-159	Sequence 159, App
18	170.8	16.0	2336	12	US-09-971-392-46	Sequence 46, Appl
19	170.8	16.0	2336	12	US-10-101-510-458	Sequence 458, App
20	167.6	15.7	1296	13	US-10-076-421-1	Sequence 1, Appli
21	167.6	15.7	2294	9	US-09-735-705-123	Sequence 123, App
22	167.6	15.7	2294	10	US-09-850-716-123	Sequence 123, App
23	167.6	15.7	2294	10	US-09-897-778-123	Sequence 123, App
24	167.6	15.7	2294	11	US-09-466-396A-123	Sequence 123, App
25	167.6	15.7	2294	12	US-10-117-982-123	Sequence 123, App
26	167.6	15.7	2344	12	US-10-101-510-383	Sequence 383, App
27	166	15.6	474	9	US-09-864-761-1682	Sequence 1682, Ap
28	166	15.6	1236	10	US-09-880-503-12	Sequence 12, Appl
29	166	15.6	2304	12	US-10-301-822-160	Sequence 160, App
30	166	15.2	2304	14	US-10-171-311-183	Sequence 183, App
31	162	15.2	2486	14	US-10-106-698-1989	Sequence 1989, Ap
32	159.8	15.0	831	10	US-09-880-503-14	Sequence 14, Appl
33	159.8	15.0	972	10	US-09-880-503-16	Sequence 16, Appl
34	139	13.1	461	11	US-09-918-995-10405	Sequence 10405, A
35	123	11.5	1959	12	US-10-101-510-743	Sequence 743, App
36	123	11.5	2035	12	US-10-101-510-448	Sequence 448, App
37	123	11.5	2048	11	US-09-858-909-1	Sequence 1, Appli
38	120.4	11.3	2036	10	US-09-880-107-1612	Sequence 1612, Ap
39	120.4	11.3	2036	10	US-09-920-300A-237	Sequence 237, App
40	114.4	10.7	116	9	US-09-864-761-18438	Sequence 18438, A
41	105	9.9	457	11	US-09-918-995-11188	Sequence 11188, A
c 42	92.4	8.7	372	10	US-09-920-300A-237	Sequence 237, App
c 43	92.4	8.7	372	13	US-10-033-528-237	Sequence 237, App
44	88	8.3	442	11	US-09-918-995-12715	Sequence 12715, A
45	83.8	7.9	1230	9	US-09-879-792-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT 1
US-09-987-455-4
; Sequence 4, Application US/0998745
; Publication No. US2003004972 9A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-4

Query Match         100.0%;  Score 1065;  DB 11;  Length 1065;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60

Qy         61 CTCACCGAGTCGGGTGCCTCCGTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
```

```
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRNT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-7

Query Match             100.0%;   Score 1065;   DB 11;  Length 1065;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1065;   Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCCACGCACAGC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCCACGCACAGC  60

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120

Qy    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTGAGAGACAGCAGTACAGCCAGCCTCAG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTGAGAGACAGCAGTACAGCCAGCCTCAG  180

Qy    181  AATCCTGATGGGGATGCCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AATCCTGATGGGGATGCCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240

Qy    241  GAGTACTGTGATGTGCCCTGCCTCCACCTGCGCGCCTGAGACAGTACAGCCAGCCTCAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GAGTACTGTGATGTGCCCTGCCTCCACCTGCGCGCCTGAGACAGTACAGCCAGCCTCAG  300

Qy    301  TTTCGCATCAAAGGAGGGGCTCTTCGCCCGACCATGCCCTCCACCCTGGCCAGGCTGCCATC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TTTCGCATCAAAGGAGGGGCTCTTCGCCCGACCATGCCCTCCACCCTGGCCAGGCTGCCATC  360

Qy    361  TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCGCCCGGGGCATACTCATCAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCGCCCGGGGCATACTCATCAGC  420

Qy    421  TCCTGCTGGATTCTCTGCGCCTCCTTCGGGGGCATGCTTCGCCCACCACCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  TCCTGCTGGATTCTCTGCGCCTCCTTCGGGGGCATGCTTCGCCCACCACCTG  480

Db    481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGGGTGCTGGGAGGAGCAGAAATTTGAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGGGTGCTGGGAGGAGCAGAAATTTGAA  540

Db    541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACACTTACGACAATTGCG  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACACTTACGACAATTGCG  600

Db    601  CTGCTGCAGCTGAAATCGGAATCGGAATCGGATTCGTCCCGTGTCCCGGAGAGCAGCGTGGTTCCGCACT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601  CTGCTGCAGCTGAAATCGGAATCGGAATCGGATTCGTCCCGTGTCCCGGAGAGCAGCGTGGTTCCGCACT  660
```

```
Qy    661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGATGTGAGCTCTCCGGC        720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGATGTGAGCTCTCCGGC        720

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  780

Qy    781 AGACTGTACCCATCCAGCCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGACTGTACCCATCCAGCCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840

Qy    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC   900

Qy    901 TGCCAGGGGCGATTCGGGAGGCCCCCTGGGCGTGTGACGAAGGATGTCCCGGGTG       960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TGCCAGGGGCGATTCGGGAGGCCCCCTGGGCGTGTGACGAAGGATGTCCCGGGTG       960

Qy    961 GGCATCATCAGCTGGGGCCTGGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACCACAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCATCATCAGCTGGGGCCTGGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACCACAAG 1020

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA                1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA                1065
```

RESULT 3
US-09-969-271-6
; Sequence 6, Application US/0996927l
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6

Query Match   99.8%;   Score 1063.4;   DB 9;   Length 1689;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 1064;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC  684

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 744

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGCCTGGGCCTGGGCAAACATAATTAATTACTGCGGG 180
          |||||||||||||||||||||||||||||||||||||||||||||
Db    745 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGCCTGGGCAAACATAATTAATTACTGCGGG 804

Qy    181 AATCCTGATGGGGATGCCCAAGCCCTGTGTTGCCCACGTGCCAGGCTGACGCTGG      240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 AATCCTGATGGGGATGCCCAAGCCCTGTGTTGCCCACGTGCCAGGCTGACGCTGG      864

Qy    241 GAGTACTGTGATGTGCCTCCTCCTTGCCTCCACCTTGCGGCCGCCTGAGACAGTACAGCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 GAGTACTGTGATGTGCCTCCTCCTTGCCTCCACCTTGCGGCCGCCTGAGACAGTACAGCAGCCTCAG 924

Qy    301 TTTCGCATCAAAGGAGGGGCTCTTGCCGACATGCCTCCACCCCTGGCAGGCTGCCATC    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    925 TTTCGCATCAAAGGAGGGGCTCTTGCCGACATGCCTCCACCCCTGGCAGGCTGCCATC    984

Qy    361 TTTGCCAAGCACACAGGAGGTCGCCCGGAGCGGTTCCTGTCGCGGGGCATACTCATCAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 TTTGCCAAGCACACAGGAGGTCGCCCGGAGCGGTTCCTGTCGCGGGGCATACTCATCAGC 1044

Qy    421 TCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTCCGCCCCACCACCTG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTCCGCCCCACCACCTG  1104

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 1164

Qy    541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  1224

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTCGCCCAGGAGGAGCAGCGGGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 CTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTCGCCCAGGAGGAGCAGCGGGTGGTCCGCACT 1284

Qy    661 GTGTGCCTTCCCCCGCGGGGACCTGCAGCTGCCGGACTGGATGTGAGCTCTCCGGC     720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 GTGTGCCTTCCCCCGCGGGGACCTGCAGCTGCCGGACTGGATGTGAGCTCTCCGGC    1344

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1404

Qy    781 AGACTGTACCCATCCAGCCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 AGACTGTACCCATCCAGCCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1464

Qy    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC  1524

Qy    901 TGCCAGGGGCGATTCGGGAGGCCCCCTGGGCGTGTCGACGAAGGATGTCCCGGGTGGTG   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 TGCCAGGGGCGATTCGGGAGGCCCCCTGGGCGTGTCGACGAAGGATGTCCCGGGTGGTG  1584

Qy    961 GGCATCATCAGCTGGGGCCTGGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACCACAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GGCATCATCAGCTGGGGCCTGGGGCGTGTGACAGAAGGATGTCCCGGGTGTGTACCACAAG 1644

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA               1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1645 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA               1689
```

RESULT 4
US-10-193-656-7
; Sequence 7, Application US/10193656
; Publication No. US2003009673lA1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X07393
; DATABASE ENTRY DATE: 1995-03-27
; RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match           99.8%;   Score 1063.4;   DB 14;  Length 2509;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 1064;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    678 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 737

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCCGTGAATCCGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    738 CTCACCGAGTCGGGTGCCTGCCTCCCGCCGCCTTCCCGTGCCTCCCGTGAATCCGATAGGCAAGGTT 797

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    798 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 857

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    858 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 917

Qy    241 GAGTACTGTGATGTGCCCTCCTGTCTGCTCACCTGCGGCTACAGCAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    918 GAGTACTGTGATGTGCCCTCCTGTCTGCTCACCTGCGGCTACAGCAGCCAGCCTCAG 977

Qy    301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCTCCCACCCCTGGCGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCTCCCACCCCTGGCGCTGCCATC 1037

Qy    361 TTTGCCAAGCACAGGAGAGTCGCCCGGCCGAGAGCCGTTCCTGTGCGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 TTTGCCAAGCACAGGAGAGTCGCCCGGCCGAGAGCCGTTCCTGTGCGGGGCATACTCATCAGC 1097

Qy    421 TCCTGCTGGATTCTCTCTGCCCGGCCGACTGCTTCCGCCCCACCCCTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TCCTGCTGGATTCTCTCTGCCCGGCCGACTGCTTCCGCCCCACCCCTG 1157

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGGTCCCTGGCGTGGTCCTGGGGCAGCAGAAACATAATCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 ACGGTGATCTTGGGCAGAACATACCGGGGTCCCTGGCGTGGTCCTGGGGCAGCAGAAACATAATCAGC 1217

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1218 GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCCG 1277

Qy    601 CTGCTGCGACGCTGAAATGGATTCGTCCCGGCAGCCGTGGTCCGCCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1278 CTGCTGCGACGCTGAAATGGATTCGTCCCGGCAGCCGTGGTCCGCCACT 1337

Qy    661 GTGTGCGTCCCCGGCCGGCAGCTGCACCTGGACTGGACGGACGTGTGAGCTTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1338 GTGTGCGTCCCCGGCCGGCAGCTGCACCTGGACTGGACGGACGTGTGAGCTTCTCCGGC 1397

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAAGGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1398 TACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAAGGCGGCTGAAGGAGGCTCATGTC 1457

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1458 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACGAACAGTCACCGAC 1517

Qy    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1518 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCC 1577

Qy    901 TGCCAGGGCGGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGCATGCCCATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1578 TGCCAGGGCGGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGCCCATGACTTTGGTG 1637

Qy    961 GGCATCATCAGCTGGGCGCTGGGCCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1638 GGCATCATCAGCTGGGCGCTGGGCCTGTGGACAGAAGGATGTCCCGGGTGTGTACACCAAG 1697

Qy   1021 GTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1698 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1742
```

```
RESULT 5
US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US2002008179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only);
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match           99.8%;   Score 1063.4;   DB 9;   Length 2519;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 1064;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    701 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 760

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCCGTGAATCCGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    761 CTCACCGAGTCGGGTGCCTGCCTCCCGTGCCTCCCGTGAATCCGATAGGCAAGGTT 820

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    821 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 880

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGTGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    881 AATCCTGATGGGGATGCCAAGCCCTGTGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 940

Qy    241 GAGTACTGTGATGTGCCCTCCTGTCTGCTCACCTGCGGCTGACAGTACAGCAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941 GAGTACTGTGATGTGCCCTCCTGTCTGCTCACCTGCGGCTGACAGTACAGCAGCCAGCCTCAG 1000

Qy    301 TTTCGCATCAAAGGAGGGGTCTTCGCCGACATCGCCTCCCACCCCTGGCGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1001 TTTCGCATCAAAGGAGGGGTCTTCGCCGACATCGCCTCCCACCCCTGGCGCAGGCTGCCATC 1060

Qy    361 TTTGCCAAGCACAGGAGAGTCGCCGGTTCGGCCATACTCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||
Db   1061 TTTGCCAAGCACAGGAGAGTCGCCGGTTCGGCCATACTCAGC 1120

Qy    421 TCCTGCTGGATTCTCTCTGCCGCCGCTTCCGCCGAGGAGAGGTTTCCGCCCACCACCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1121 TCCTGCTGGATTCTCTCTGCCGCCGCTTCCGCCGAGGAGAGGTTTCCGCCCACCACCTG 1180

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCGGGGAGAGAGGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 1181 ACGGTGATCTTGGGCAGAACATACCGGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 1240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600

Db 1241 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGCTGTGTGCCGGACTGGACGTGGTTCCGCACT  660

Db 1301 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGCTGTGTGCCGGACTGGACGTGGTTCCGCACT 1360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACCTGCCGGACGTGGTGAGCTCTCCGGC  720

Db 1361 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACCTGCCGGACGTGGTGAGCTCTCCGGC 1420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCGATTCGGAGCGGCGGCTGACCTGACAAGCCTCATGTC  780

Db 1421 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCGATTCGGAGCGGCGGCTGACCTGACAAGCCTCATGTC 1480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTACTTAACAGAACAGTCACCGAC  840

Db 1481 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTACTTAACAGAACAGTCACCGAC 1540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCCAGCGCAAACTTGCACGACGCC  900

Db 1541 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCCAGCGCAAACTTGCACGACGCC 1600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901 TGCCAGGGCGATTCGGGAGGCCCTGGGGCTGTGTGGACAGAAGAGGATGTCCCGGGTGTGTACACAAAG  960

Db 1601 TGCCAGGGCGATTCGGGAGGCCCTGGGGCTGTGTGGACAGAAGAGGATGTCCCGGGTGTGTACACAAAG 1660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961 GGCATCATCAGCTGGGGGCCTGGGCTGTGCTGGAGACACTTGCGAGACGCC 1020

Db 1661 GGCATCATCAGCTGGGGGCCTGGGCTGTGCTGGAGACACTTGCGAGACGCC 1720
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065

Db 1721 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1765
```

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCACCGTGCCACCGGACACAGC   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCACCGTGCCACCGGACACAGC  126

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  186

Qy  121 TACACAGCAGACAGACCCCAGTGCCCAGCACTGGGCCTGGCAAACATAATTACTGCCGG  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  187 TACACAGCAGACAGACCCCAGTGCCCAGCACTGGGCCTGGCAAACATAATTACTGCCGG  246

Qy  181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  247 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  306

Qy  241 GAGTACTGTGATGTGCCCCTGCTGCCACCTGCCGAGACAGTACAGCCAGCCTCAG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  307 GAGTACTGTGATGTGCCCCTGCTGCCACCTGCCGAGACAGTACAGCCAGCCTCAG  366

Qy  301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCGCCTGGCAGGCTGCCATC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  367 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCGCCTGGCAGGCTGCCATC  426

Qy  361 TTTGGCCAAGCACAGGAGGTCGCCGGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  427 TTTGGCCAAGCACAGGAGGTCGCCGGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC  486

Qy  421 TCCTGCTGGATTCTTCTCTGCCGCCACTGCTTCCGCCCCACCACCTG  480
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  487 TCCTGCTGGATTCTTCTCTGCCGCCACTGCTTCCGCCCCACCACCTG  546

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAA  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  547 ACGGTGATCTTGGGCAGAACATACCGGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAA  606

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  666

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGCTGTGTGCCGGACTGGACGTGGTTCCGCACT  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  667 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGCTGTGTGCCGGACTGGACGTGGTTCCGCACT  726

Qy  661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACCTGCCGGACGTGGTGAGCTCTCCGGC  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  727 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACCTGCCGGACGTGGTGAGCTCTCCGGC  786

Qy  721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGACCTGACAAGCCTCATGTC  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  787 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGACCTGACAAGCCTCATGTC  846

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTACTTAACAGAACAGTCACCGAC  840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  847 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTACTTAACAGAACAGTCACCGAC  906

Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCCAGCGCAAACTTGCACGACGCC  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  907 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCCAGCGCAAACTTGCACGACGCC  966

Qy  901 TGCCAGGGCGATTCGGGAGGCCCTGGTGTCGTGCAGGATGCGGCTGTGAACGATGCCGATGACTTTGGTG  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  967 TGCCAGGGCGATTCGGGAGGCCCTGGTGTCGTGCAGGATGCGGCTGTGAACGATGCCGATGACTTTGGTG 1026

Qy  961 GGCATCATCAGCTGGGGCCGTGTGCTGACAGAAGAGATGTGTCCCGGGTGTGTACACAAAG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1027 GGCATCATCAGCTGGGGCCGTGTGCTGACAGAAGAGATGTGTCCCGGGTGTGTACACAAAG 1086

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1062
        ||||||||||||||||||||||||||||||||||||||||||
Db 1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

Query Match   99.7%; Score 1062; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1062; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

RESULT 7
US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match   99.7%; Score 1062; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1062; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 TCTGAGGGAAACAAGTGACTGCTACTTTGGGAATGGTCAGCCTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67 TCTGAGGGAAACAAGTGACTGCTACTTTGGGAATGGTCAGCCTGGCACGCACAGC 126

Qy    61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 186

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 246

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGAACCCGCAGCGTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGAACCCGCAGCGTGACGTGG 306

Qy   241 GAGTACTGTGATGTGCCCCTCGTCTCCACCTGCGGCCCTGAGACAGCAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 GAGTACTGTGATGTGCCCCTCGTCTCCACCTGCGGCCCTGAGACAGCAGCCAGCCTCAG 366

Qy   301 TTTCGCATCAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 TTTCGCATCAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGCTGCCATC 426

Qy   361 TTTGCCAAGCACAGGAGGAGTCGCCCGAGCAGCCGGTTCCTGTGCGGGGAGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 TTTGCCAAGCACAGGAGGAGTCGCCCGAGCAGCCGGTTCCTGTGCGGGGAGCATACTCATCAGC 486

Qy   421 TCCTGCTGGATTCTCTGCCGCCGCCACTGGTTCCGCCCCCACCACCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 TCCTGCTGGATTCTCTGCCGCCGCCACTGGTTCCGCCCCCACCACCTG 546

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCTGCGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 ACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCTGCGAGGAGCAGAAATTTGAA 606

Qy   541 GTCGAAAATACATTGTCCATAAGGAATTGTGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GTCGAAAATACATTGTCCATAAGGAATTGTGATGATGACACTTACGACAATGACATTGCG 666

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGCAGCGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGCAGCGTGGTCCGCACT 726

Qy   661 GTGTGCCTTCCCCGGCGACCTGCAGTCGCCGACTGACGTGAGCTGTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 GTGTGCCTTCCCCGGCGACCTGCAGTCGCCGACTGACGTGAGCTGTCCGGC 786

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGACGGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGACGGCGGCTGAAGGAGGCTCATGTC 846

Qy   781 AGACTGTACCCATCCAGCCGCTGCACATCACAACACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 AGACTGTACCCATCCAGCCGCTGCACATCACAACACATTTACTTAACAGAACAGTCACCGAC 906

Qy   841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACTTGCACGACGCC 966

Qy   901 TGCCAGGGCGGATTCGGGGAGGCCCCCTGGTGTGTCTGACGATGGCCGACGACTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 TGCCAGGGCGGATTCGGGGAGGCCCCCTGGTGTGTCTGACGATGGCCGACGACTTTGGTG 1026

Qy   961 GGCATCATCAGCTGGGCCTGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1027 GGCATCATCAGCTGGGCCTGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1086

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1062
          ||||||||||||||||||||||||||||||||||||||||||
Db  1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

RESULT 8
US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US2002015623A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
;
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

Query Match   99.7%; Score 1061.8; DB 10; Length 2641;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 1063; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy     1 TCTGAGGGAAACAAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   821 TCTGAGGGAAACAAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 880

Qy    61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   881 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 940
```

```
Qy    121 TACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941 TACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 1000

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTCGCCACGTGCCAGGCCTGGTGAAGAACCGCAGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1001 AATCCTGATGGGATGCCAAGCCCTGGTCGCCACGTGCCAGGCCTGGTGAAGAACCGCAGG 1060

Qy    241 GAGTACTGTGATGTGCCCTCCTGTCCACCTGCGGCCTGAGACAGTAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1061 GAGTACTGTGATGTGCCCTCCTGTCCACCTGCGGCCTGAGACAGTAGCCAGCCTCAG 1120

Qy    301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1121 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 1180

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCGGCCCCACTGTTCCGCCCACACCTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1181 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCGTGCGGGGGGCATACTCATCAGC 1240

Qy    421 TCCTGCTGGATTCTCTGCGCCCACTGCTTCCGCCCCACACCTG 480
          |||||||||||||||||||||||||||||||||||||||||||
Db   1241 TCCTGCTGGATTCTCTGCGCCCACTGCTTCCGCCCCACCACCTG 1300

Qy    481 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGCCGGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1301 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGCCGGAGGAGCAGAAATTTGAA 1360

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1361 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1420

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1421 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 1480

Qy    661 GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGACTGCCGAGTGTGAGCTTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1481 GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGACTGCCGAGTGTGAGCTTCTCCGGC 1540

Qy    721 TACGGCAAGCAAGCATGAGGCCTTGTCCTTTCATTCGGAGCGGCGCTGAAGGAGGCTCATGTC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1541 TACGGCAAGCAAGCATGAGGCCTTGTCCTTTCATTCGGAGCGGCGCTGAAGGAGGCTCATGTC 1600

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1601 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1660

Qy    841 AACATGCGTGTGCTGAGACACTCGGAGCGGCGGGCGGGCCCCGTGTGTCTGAACGATGGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1661 AACATGCGTGTGCTGAGACACTCGGAGCGGCGGGCGGGCCCCGTGTGTCTGAACGATGGCG 1720

Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCCATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1721 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCCATGACTTTGGTG 1780

Qy    961 GGCATCATCAGCTGGGCCTGGGCGTGTGTGGACAGAAGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1781 GGCATCATCAGCTGGGCCTGGGCGTGTGTGGACAGAAGATGTCCCGGGTGTGTACACCAAG 1840

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACAACATGCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1841 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACAACATGCGACCGTGA 1885
```

```
APPLICANT: NGUYEN, CATHERINE
APPLICANT: VIENS, PATRICE
APPLICANT: FERT, VINCENT
TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
FILE REFERENCE: 1546-R-00
CURRENT APPLICATION NUMBER: US/10/007,926A
CURRENT FILING DATE: 2001-12-07
PRIOR APPLICATION NUMBER: 60/254,090
PRIOR FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 468
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 433
LENGTH: 329
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)
FEATURE:
NAME/KEY: modified_base
LOCATION: (57)
OTHER INFORMATION: a, t, c or g
FEATURE:
NAME/KEY: modified_base
LOCATION: (82)
OTHER INFORMATION: a, t, c or g
FEATURE:
NAME/KEY: modified_base
LOCATION: (159)
OTHER INFORMATION: a, t, c or g
FEATURE:
NAME/KEY: modified_base
LOCATION: (264)
OTHER INFORMATION: a, t, c or g
FEATURE:
NAME/KEY: modified_base
LOCATION: (278)
OTHER INFORMATION: a, t, c or g
US-10-007-926A-433

Query Match        21.6%;  Score  230.2;  DB 12;  Length 329;
Best Local Similarity  97.5%;  Pred. No. 4.3e-63;
Matches  232;  Conservative   0;  Mismatches   6;  Indels   0;  Gaps   0;

Qy    828 AACAGTCACCGACAACATGCTGTGTCTGGAGCACTCGGAGCGCGGGCCCCAGGCAAA 887
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1 AACAGTCACCGACAACATGCTGTGTCTGGAGCACTCGGAGCGCGGGCCCCAGGNAAA 60

Qy    888 CTTGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCG 947
          ||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db     61 CTTGCACGACGCCTGCCAGGGNATTCGGGNATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCG 120

Qy    948 CATGACTTTGGTGGGCATCATCAGCTGGGACTGGGACAGAAGGATGTCCCGGG 1007
          |||||||||||||||||||||||||||||||||||       ||||||||||||
Db    121 CATGACTTTGGTGGGCATCATCAGCTGGGACTGGGACNTNGGACAGAAGGATGTCCCGGG 180

Qy   1008 TGTGTACACAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 238

RESULT 10
US-09-880-503-15
; Sequence 15, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
```

RESULT 9
US-10-007-926A-433
; Sequence 433, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL

```
Qy    897 CGGCTGCCAGGGCGGAGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTT  956
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 TTCCTGCCAGGGAGACTCAGGGGGACCCCGTCGTCTTCCCTCCAAGGCCGCATGACTTT 1085

Qy    957 GGTGGGCATCATCAGCTGGGCCTGGGCTGTGGGACAGAGGATGTCCCGGGTGTGTACAC 1016
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 GACTGGAATTGTGAGCTGGGCCGTGAGTGTGCCCTGAAGGAGCAAGCCAGGCGTCTACAC 1145

Qy   1017 AAAGGTTACCAACTACCTAGACTGGATTCG 1046
          ||| |  ||| |||||||||| ||||||||
Db   1146 GAGAGTCTCCACACTTCTTACCCTGGATCCG 1175
```

RESULT 11
US-10-198-846-12748
; Sequence 12748, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12748
; LENGTH: 1415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 1414, 1415
; OTHER INFORMATION: n = A, T, C or G
US-10-198-846-12748

Query Match   17.4%; Score 185; DB 14; Length 1415;
Best Local Similarity   92.6%; Pred. No. 2e-48;
Matches 249; Conservative   0; Mismatches 15; Indels   5; Gaps   5;

Qy   1 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCCACGCACAGC   60
          ||||||||||||||||||||| |||| |||||||||||||||||||||||||||||||
Db   906 TCTGAGGGAAAACAGTGACTG-TACTTTGGGAATGGGTCAGCCTACCGTGCCACGCACAG- 963

Qy   61 CTCACGGAGTCGGGTGCCTCCT-GCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGT 119
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   964 CTCACGGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGT 1023

Qy   120 TTACACAGCACAGAACCCCAGTGCCACCCCAGTGCCAGGCACTGCTGGGCTGGGCAAACATAATTACTGCCG 179
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 TTACACAGCACAGAACCCCAGTGCCACCCCAGTGCCAGGCACTGCTGGGCTGGGCAAACATAAATTACTGCC 1083

Qy   180 GAATCCTGATGGGGATGGCAAGCCCTGCTGCCACCTGCTGAA-GAACCGCAGGCTGACGT 238
          ||| |||||||||||| || |||||||| |||||||||||||| || || ||||||||||
Db   1084 GAAT-CTGATGGGGATGGCCAAGCCCTGTTTGCACGTGCTGCGCAGCAGCCAGGCTGACGT 1142

Qy   239 GGGAGTACTGTGATGTGCCCTCCTCC 267
          ||||||||||||||||||||||| ||
Db   1143 GGGAGTACTGTGATGTGGCCTCTGTCACC 1171
```

RESULT 12
US-09-735-705-122
; Sequence 122, Application US/09735705
; Patent No. US20020523229A1
; GENERAL INFORMATION:

---

; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;  LENGTH: 1212
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-880-503-15

Query Match      18.0%;  Score 191.6;  DB 10;  Length 1212;
Best Local Similarity    52.3%;  Pred. No. 1.5e-50;
Matches  549;  Conservative   0;  Mismatches  459;  Indels  42;  Gaps   4;

```
Qy     18 CTGCTACTTTGGGAATGGGTCAGCCTACCGTGCCACGCACCAGCCTCACCGAGTCGGGTGC   77
          ||||||  ||||||||||                ||||
Db    147 CTGCTATGAGGGGAATGGTCACTTTACCGAGGAAAAGGCCAGCACTGACACCATGGGCCG  206

Qy     78 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACAGCACAGAACCC  137
          |||||  ||   ||   | |||
Db    207 GCCCTGCCTGCCACTCTGCCACTGTCCTTCAGCAAACGTACCACCATGCCCACAGATC  266

Qy    138 CAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC  197
          |||   |   ||   ||| |||||| ||||| |  ||| |||  | |  |||  ||||||
Db    267 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCAGGAACCCAGACACACCGGAG  326

Qy    198 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC  257
          |||   ||  ||||   |||  ||   | ||| ||||| |||||||||||||||||||||
Db    327 GCGACCCTGGTATGTGCAGTGGGCCTAAAGCCCGCTTGTCCAAGAGTGCATGGTGCA  386

Qy    258 CTCCTGCTCCACCTGCGGCCTGAGACAGTAGCAGCCA--------GCCTCAGTT  302
          |||||||||       |||||| |||||||| ||||||        ||||||
Db    387 TGACTGCGCCAGATGGAAAATTAAAATTCAGTGTGGCCAAAAGACTCTGAGGCCCCCGCTT  446

Qy    303 TCGGCATCAAAGAGAGGGGCTTCTCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATCTT  362
          |||       |||||| |  ||||||| ||
Db    447 TAAGATTATTGGGGGAGAATTCACCACCATCGAGAACCAGCCCCTGGTTGCCGCCATCTA  506

Qy    363 TGCCAAGCACAGGAGGAGTCGCCCCGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGCTC  422
          |||   |    ||   |||       || |       |||  ||| |||||||||||||
Db    507 CAGGAGGCACCGGGGGGGGCTCTGTCA---CCTACGTGTGTGGAGGCAGCCTCATCAGCCC  563

Qy    423 CTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCTGAC  482
          ||  ||     || |      |||  || |  ||||  || ||||||
Db    564 TTGCTGGGTGATCAGCGCCACACACCGCTTCATTGATTACCCAAGAAGAGGAGGACTACAT  623

Qy    483 GGTGATCTTGGGCAGAACATACCGGGTGGTCCCCTGGCGAGGAGGAGCAGAAATTTGAAGT  542
          | | |||  |||||   ||  ||||||| ||  ||| ||  |  ||| ||| | || ||||
Db    624 CGTCTACCTGGGTCGCTCAAGGCTTAACTCCAACACGCAAGGGGAGATGAAGTTTGAGGT  683

Qy    543 CGAAAAATACATTGTCCATAAGGAAATTCGATGATGACA------CTTACGACAATGACAT  596
          ||||   |    |  || |   ||  ||   || ||  ||      |  || || || |||
Db    684 GGAAAACCTCATCCTACACAAGGACTACACAGCGCTTGCTCCACCAAAGACGACTTATACA  743

Qy    597 TGGCCTGCTGCAGCTGAAATCGGATTCGTCCCCGTGTGCCAGGAGAGCAGCCGTGGTCCG  656
          ||  |||  || ||   ||  |  ||  | |   |||   |||  || ||||| |||  ||
Db    744 TGCCTTGCTGAAGATCCGTTCCAAGAGGGGCAGCCATCCGGACTATACA  803

Qy    657 CACTGTGTCCTTCTCCCCGGCGGACCTGCAGCTGCCAGCTGCGACGGAGTGTGAGCTCTC  716
          |||  |  | | | |  |||| |  ||||| |||  ||  || | || ||  |||   ||
Db    804 GACCATCTGCCTGCCCTGCCCTGCGATGTATAACGATCCCCAGTTTGGCACAACGTGTGAGATCAC  863

Qy    717 CGGCTACGCAAGCATGAGGCGGCGCGCTGAAGGACGCGGCTGAAGGAGGAGGCTCA  776
          ||      ||    |  || ||||| ||         | || || ||||||||||||
Db    864 TGGCTTTCGAAAAGAGAATTCTACCGACTATCTCTATCCGAGCAGCAGTGAAAATGACTGT  923

Qy    777 TGTCAGACTGTACCCATCCAGCGCCGGCTGCACATCACAGAACAGTCAC  836
          ||  ||    ||      ||       | |  || |  ||  | | ||||
Db    924 TGTGAAGCTGATTTCCCACTTCCCGAGCAGTGTGTGGAGACACTCGGAAGTCAC  983

Qy    837 CGACAAACATGCTGTGTGCTGGAGCACTCGGAGCCGGCGGCCCCCAGGCAAACTTGCACGA  896
          ||     |  |  ||||||||||||                            ||||  ||
Db    984 CACCAAAATGCTATGTGTGCTGACCCCC---------------------AATGGAAAACAGA 1025
```

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
; , ORGANISM: Homo sapien
US-09-735-705-122

Query Match     16.0%;    Score 170.8;    DB 9;    Length 1475;
Best Local Similarity   51.3%;    Pred. No. 7e-44;
Matches  551;    Conservative  0;    Mismatches 457;    Indels 66;    Gaps 4;

Qy    18 CTGCTACTTTGGGAATGGTCAGCCTACCGTGCCACGCACCAGCCTCACCGAGTCGGGTGC 77
Db   287 CTGCTATGAGGGGAATGGTCACTTTTACCGGAGGAAAGCCACGACACTGACACCATGGGCCG 346

Qy    78 CTCCTGCCTCCCGTGGAATCCATGATCCTGATAGGCAAGGTTTACACAGCAGACAGAACCC 137
Db   347 GCCCTGCCTGCCCTGGAACTCTGCCACTGTCCTGTCCTTCAGCAAACGTACCACACAGATC 406

Qy   138 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCCGGAATCCTGATGGGGATGC 197
Db   407 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCGAGGAACCCAGACAACCGGAG 466

Qy   198 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC 257
Db   467 GCGAACCCTGGTGCTATGTGCAGTGGGCCTAAAGCCCGCTTGTCCAAGAGTGCATGGTGCA 526

Qy   258 CTCCTGCTCCACCTGCGGCGCCTGAG------------------------------- 281
Db   527 TGACTGCGCAGATGGAAAAAAGCCCTCTCTCCAGAAGAATTAAAATTTCAGTGTGG 586

Qy   282 ---ACAGTACAGCCAGCCCAGTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTC 338
Db   587 CCAAAGAGACTCTGAGGCCCCGCTTTAAGATTATTGGGGGAGAATTCACCACCATCGAGAA 646

Qy   339 CCACCCCTGGCAGGCTGCCATCTTGCCAAGACAGCAGAGGTCGCCCGAGAGCGGGTTCCT 398
Db   647 CCAGCCTGGTTTGCGGGTTTGCGGCGTCTGTCTGTCACCTACGT 703

Qy   399 GTGCGGGGGCATACTACTCATCAGCTGCCTGGAATTCCTTCTGCGCGCCCCACTGCTTCCAGGA 458
Db   704 GTGTGGAGGCAGCGCCATCAGCGCCGTTGTCGGGTGATCAGCCGCCACACTGCTTCATTGA 763

Qy   459 GAGGTTTCCGCCCCCACCACCTGACGGTGATCTTGGGCAGAGAACAACATACCGGGGTGGTCCCTGG 518
Db   764 TTACCCAAGAAGGAGGAGCACTACATCGTCTACCTGGGTCGCTCAAGGCTTAACTCCAAACAC 823

Qy   519 CGAGGAGGAGCAGAAATTTGAGTCGAAAAATACATTGTCCATAAGAATTCGATGATGA 578
Db   824 GCAAGGGGAGATGAAGTTTGAGGTGGAAAAACCTCATCCTACACAAGGACTACAGCCGCTGA 883

Qy   579 CA------CTTACGACAATGACAATTGCGCGTGCTGCAGCTGAAATCGGATTCGTCCCCCTG 632
Db   884 CACGCGTTGCTCCACCACGACCATTGCCTTGCTGCTATGTGCAGGTGGCGCATGGTGGTG 943

Qy   633 TGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCGGCGGCGGACCTCAGCTGCC 692
Db   944 TGGCAGCCATCCGGACTATACAGACCATCTGCCTGCCCTGCCTTCGATGTATAACGATCCCCA 1003

Qy   693 GGACTGGACGGAGTGTGAGCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTA 752
Db  1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy   753 TTCGGAGCGGCTGAAGGAGGCTCATGTCAGACGTGACGTGTACCATCCAGCCGCGTGCACATCACA 812
Db  1064 TCCGGAGCTGCTGAAGAATGATGACTGTTGTGAACGCAGACCGGGAGTGTGCAGCAGCC 1123

Qy   813 ACATTTACTTAACAGACAGCAGTGCTCACCGCACAACATGCTGTGTGCTGGAGACACTCGGAGCGG 872
Db  1124 CCACTACTACGGCTCTGAAGTCACCACCACCAAAATGCTGTGTGCTGCTGACCCACAGTG--- 1180

Qy   873 CGGGCCCAGGCAAACTTGCACGACGCGCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTG 932
Db  1181 --------GAAAACAGATTCCTGCCAGGAGACTCAGGGGGACCCCTCGTCTG 1225

Qy   933 TCTGAACGATGGCCGCATGACTTTGGTGGCAGTCATCATCAGCTGGGGCTGTGTGGACA 992
Db  1226 TTCCCTCCAAGGCGCGCATGACTTTGACTGGAATTGTGAGCTGGGAGCTGGGGCGTG 1285

Qy   993 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATTCG 1046
Db  1286 GAGGACAAGCCAGGCGGTCTACACGACGGGTCTCACACTTCTTACCCTGGATCCG 1339
```

```
RESULT 13
US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US2002011513A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
; , ORGANISM: Homo sapien
US-09-850-716A-122

Query Match     16.0%;    Score 170.8;    DB 10;    Length 1475;
Best Local Similarity   51.3%;    Pred. No. 7e-44;
Matches  551;    Conservative  0;    Mismatches 457;    Indels 66;    Gaps 4;

Qy    18 CTGCTACTTTGGGAATGGTCAGCCTACCGTGCCACGCACCAGCCTCACCGAGTCGGGTGC 77
Db   287 CTGCTATGAGGGGAATGGTCACTTTTACCGGAGGAAAGCCACGACACTGACACCATGGGCCG 346

Qy    78 CTCCTGCCTCCCGTGGAATCCATGATCCTGATAGGCAAGGTTTACACAGCAGACAGAACCC 137
Db   287 CTGCTATGAGGGGAATGGTCACTTTTACCGGAGGAAAGCCAGCACTGACCACCACGGGCCG 346

Qy    78 CTCCTGCCTCCCGTGGAATCCATGATCCTGATAGGCAAGGTTTACACAGCAGACAGAACCC 137
Db   347 GCCCTGCCTGCCCTGGAACTCTGCCACTGTCCTGTCCTTCAGCAAACGTACCACACAGATC 406

Qy   138 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCCGGAATCCTGATGGGGATGC 197
Db   407 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCGAGGAACCCAGACAACCGGAG 466

Qy   198 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC 257
Db   467 GCGAACCCTGGTGCTATGTGCAGTGGGCCTAAAGCCCGCTTGTCCAAGAGTGCATGGTGCA 526

Qy   258 CTCCTGCTCCACCTGCGGCGCCTGAG--------- 281
Db   527 TGACTGCGCAGATGGAAAAAAGCCCTCTCTCCAGAAGAATTAAAATTTCAGTGTGG 586
```

```
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-122

Query Match         16.0%;    Score 170.8;   DB 10;   Length 1475;
Best Local Similarity   51.3%;   Pred. No. 7e-44;
Matches   551;   Conservative   0;   Mismatches 457;   Indels  66;   Gaps   4;
```

```
Qy    18  CTGCTACTTGGAATGGTCAGCTCAGCGCACGCACCTGCAGAGCTCACCGAGTCGGGTGC   77
              ||||      |||||||||||        ||||       |||   |  |
Db   287  CTGCTATGAGGGGAATGGTCACTTTACCGGAGAAAGCCAGCACTGACCATGGGCCG    346

Qy    78  CTCCTGCCTCCCGTGGAATTCCATGATCCATGATCCTCGATAGGCAAGGTTTACACAGCACAGAACCC   137
              |||||||| |||||||||| ||    ||  ||  | || |||      |||   ||||
Db   347  GCCCTGCCTGCCCTGGAACTCTGCCACTGTCCCACTGTCCCTCAGCAAACGTACCATGCCCACCAGATC   406

Qy   138  CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTACTGCCGAATCCTGATGGGGATGC   197
              |||  |||  | |||  ||| ||| ||| |||| ||||||| | |  |   |||||||||
Db   407  TGATGCTCTTCAGCTGCGGCCTGGGAAACATAATTACTACTGCGAGGAACCCAGACAACCGGAG   466

Qy   198  CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC   257
              ||||||| |||  | ||||| ||   |  || |||  | | ||| ||||   ||| |||
Db   467  GCCACCCTGGTGTCATGTGCAGGTGGGCCTAAAGCCCGTTGTCCAAGACGTGCATGGTGCA   526

Qy   258  CTCCTGCTCCACCTGCGGCCTGAG-------------   281
              |||
Db   527  TGACTGCGCAGATGGAAAAAGCCCTCTCCTCCTCCAGAAGAATTAAAATTCAGTGTGTG    586

Qy   282  ---ACAGTACAGCCAGCTCAGTTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTC   338
                 |  |||||  |||  ||||| |||||||  |  |||||| |||||   ||||||||
Db   587  CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAGAATTCACCACCATCGAGAA    646

Qy   339  CCACCCCTGGCCAGCTGCCATCTTTGCCAAGCACAGAGAGTCCCCGGAGAGCGGTTCCT   398
              || |||||| |    |||||||  |   |  || | ||||  |||||||||||  |||
Db   647  CCAGCCCTGGTTTGCGGCCATCAGAGGAGGGCACCGGGG---GGGCTCTGTCACCTACGT    703

Qy   399  GTGCGGGGGCATACTCATCAGCTCCTGCTGGGATTCTCTGCCGCCCACCTGCTTCCAGGA   458
              ||| |  ||||  ||||||||   ||||||||| |||| |||||| || |||||||| |||
Db   704  GTGTGGAGGCAGCCTCATCAGCCCTTGCTGGGTGATCAGCCGCCACACACTGCTTCATTGA    763

Qy   459  GAGGTTTCCGCCCCCACCACCTGACGCGTGATCTTGGGCAGAACATACCGGGTGTCCCTGG   518
              |||||||  ||||  || ||||| |   | | |||||||||||| || |  ||| |||
Db   764  TTACCCAAAGAAGGAGGACTACATCGTCTACCTGGGTCGCTCAAGCGTTAACTCCAACAC    823

Qy   519  CGAGGAGGAGCAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGA   578
              | |||||| |||| | ||| | ||| | |||||   |||  | |||||| ||  ||||  ||
Db   824  GCAAGGGGAGATGAAGTTTGAGGTGGAAAAACCTCATCCTCACACAAGGAACTACAGCGCTGA   883

Qy   579  CA------CTTACGACAATGACATTGCGCTGCTGCAGCTGAAGATCGGTTCCAAGGAGGCAGGTG   632
              ||       |||  || | |||||||| |||||||||       ||||||||  |   |||||
Db   633  TGCCCAGGAGAGCAGCCGTGGTCCGACTGGTCCGACTATACAGGCGGTTAATGTGTCCAGCTGCC   692

Qy   633  TGCCCAGGAGAGCAGCCGTGGTCCGACTGGTCCGACTATACAGGCGGTTAATGTGTCCAGCTGCC   692
              ||| |||||||||||||||||| ||||||| ||  | ||  |||||  | | | |||||||||
Db   944  TGGCAGCCATCCCGGACTATACAGGAGGATTCTACCGACTATCCCA   1003

Qy   693  GGACTGGACGGAGTGTGAGCTCTTCCGGCTCTGTGAGCTCCTTGTCTGTCTCCTTTCTA   752
              || |||| | ||||||||    || ||  |||  |  | |   ||  |||  ||||| |||
Db  1004  GTTTGGCACAAGCTGTGAGATCACTGAGGAGGGCTTTGGTGGGGAGAATAGAGAATTCTACCGACTATCTCTA   1063

Qy   753  TTCGGACCATGGCCGCATGAGTTGGTGGGCGCGTCATCAGCTGGGCATCATCAGCTGGGCATCATCAGC   812
              |||||| ||| ||  ||  |||| | ||||  |  | ||||||||  | |||||
Db   823  GGACTGGACGGAGTGTGAGCTCTTCCGGCTCTGTGAGCTCCTTGTCTGTCTCGCACATCACA   812

Qy  1064  ACATTTACTTAACAGAACAGTCACCGACAACACATGCTGTGTGGACACTCGGAGGGG   1123
              | | |||||  |||| | |||   ||| |||| |||| | |  | |  |||| ||  |
Db  1124  CCACTACACGGCTCTGAAGTCACCACCAAAATGCTGTGTGCTGACCCCAAGTG---   1180

Qy   813  ACATTTACTTAACAGAACAGTCACCGACAACACATGCTGTGTGGACACTCGGAGGGG   872
              |||  ||| || |  |  || || |   ||||   ||  | | ||||||||||||||
Db  1286  GAAGGACAAGCCAGGCCGCGCTCTACACGACGCGTCTACACACTTCTTACCCCTGGATCCG   1339

Qy   873  CGGGCCCCAGGCAAACTTGCAGACGCGCTGCCAGGCGGCGATTCGGAGGCGCCCCTGGTGTG   932
```

RESULT 14
US-09-897-778-122

```
; Sequence 122, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
;   APPLICANT:  Wang, Tongtong
;   APPLICANT:  Marnerakis, Margarita
;   APPLICANT:  Fanger, Gary R.
;   APPLICANT:  Vedvick, Thomas S.
;   APPLICANT:  Carter, Darrick
;   APPLICANT:  Watanabe, Yoshihiro
;   APPLICANT:  Henderson, Robert A.
;   APPLICANT:  Peckham, David W.
;   APPLICANT:  Fanger, Neil
;   TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION:   AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C16
;   CURRENT APPLICATION NUMBER: US/09/897,778
;   CURRENT FILING DATE: 2001-06-28
```

```
Db   1181 --------------GAAAACAGATTCCTGCCAGGGAGGACTCAGGGGGACCCCTCGTCTG 1225
                          ||||||||||  ||  ||||    |||||  |||||||||||
Qy    933 TCTGAACGATGGCCGCATGACTTGGTTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACA 992
            ||||||||||| |||||   ||| ||  ||||  |||   |||   |
Db   1226 TTCCCTCCAAGGCCGCATGACTTGACTGGAATTGTGAGCTGGGCCGTGGGATGTGCCCT 1285

Qy    993 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATTCG 1046
            ||||||   || ||||  || |||||||||||||  |  || |||||||||||
Db   1286 GAAGGACAAGCCAGGCGTCTCACGAGTCTCACACTTCTTACCCTGGATCCG 1339
```

RESULT 15
US-09-466-396A-122

; Sequence 122, Application US/09466396A
; Publication No. US2003011976 3A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-122

Query Match   16.0%; Score 170.8; DB 11; Length 1475;
Best Local Similarity 51.3%; Pred. No. 7e-44;
Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

```
Qy    18 CTGCTACTTTGGGAATGGTCAGCTTACCGTGCCACGCCACAGCCTCACCGAGTCGGGTGC 77
           |||||  ||||||   |||||||||   ||||||||||||  |||||||||||||  ||
Db   287 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAGCGCCACTGACACCATGGGCCG 346

Qy    78 CTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTTTACACAGCACAGAACCC 137
           ||||||    ||||||||  ||||||  |||||   || |  ||||||| ||||| ||
Db   347 GCCCTGCCTGCCCTGGAACTCTGCCACTCTGCCTCCTTCAGCAAACGTACCACCATGCCCACAGATC 406

Qy   138 CAGTGCCCAGGCACTCGGGCCTGGGCAAACATAATTACTACCGGAATCCTGATGGGGATGC 197
           ||||||||||||||||||||   |||||||  ||||||||||||||||||||||||
Db   407 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTACCGAGGAACCCAGACAACCGGAG 466

Qy   198 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC 257
           ||||||||||||   |||   ||||||||||||||| ||||||||||||| ||||||
Db   467 GCCGACCCCTGGTGCTATGTGCAGGTGGCCTAAAGCCCGTTGTCCAAGAGTGCATGGTGCA 526

Qy   258 CTCCTGCTCCACCTGCGGCCTGAG------------------------ 281
           ||||||||||||||||||||||||
Db   527 TGACTGCGCAGATGGAAAAAGCCCCTCCTCTCCAGAAGAATTAAAATTTCAGTGTGG 586

Qy   282 ---ACAGTACAGCCAGCCCAGTTCCAGTTCGGCATCAAAGGAGGGGCTTCTTGCCGACATCGCCTC 338
              ||||||||||||||||||||||||||  ||||||| |||||||| | ||||||||||||
Db   587 CCAAAAGACTCTGAGGCCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGAGAA 646

Qy   339 CCACCCCTGGCAGGCTGCTGCATCTTTGCCAAGCACAGAGGAGGTCGCCCGAGAGCGGTTCCT 398
           ||||||||||||   ||||||  |||| |||  ||||||||||
Db   647 CCAGCCCTGGTTTGCGGCCATCTACAGGAGGCACACCGGGG---GGGCTCTGTGTCACCTACGT 703

Qy   399 GTGCGGGGGCATACTCATCAGCTCCTGCTGGATTCTTCTGCGCGCCCACTGCTTCCAGGA 458
           |||||||||| |||||||||||||||||||||||||||||||
Db   704 GTGTGGAGGCAGCCTCATCAGCCCTTGCTGGGTGATCAGCCGCACACACTGCTTCATTGA 763

Qy   459 GAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGG 518
           | || ||||||||||||||||||   |||||||||||
Db   764 TTACCCAAAGAAGGAGGACTACATCGTCTACCTGGGTCGCCTCAAGGCTTAACTCCAACAC 823

Qy   519 CGAGGAGGACAGAAATTTGAGTGCAGAAATACATTGTCCATAAGGAATTCGATGATGA 578
           | |||  ||   ||   |||||| |  |
```

```
Db   824 GCAAGGGGGAGATGAAGTTTGAGGTGGAAAAACCTCATCCTACACAAGGACTACAGCGCTGA 883
                   ||||||  ||||   ||    |||||
Qy   579 CA-------CTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCGTG 632
           ||       |||||||  |||||||  ||||
Db   884 CACGCTTGCTCACCACACGACATTGCCGTTCCAAGGATCCGTTCCAAGGAGGGCAGGTG 943

Qy   633 TGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTCCTGTCCTTCCCCGGCGGACCTGCAGCTGCC 692
           ||||||||||||||  |||||||||||||   ||
Db   944 TGGGCAGCCATCCGGACTATACAGACACCATCTGCCTGCCCTGATGTATAACGATCCCCA 1003

Qy   693 GGACTGGACGGAGTGTGAGCTCCGGCAAGCATGAGCGCCTTGTCTCCTTTCTA 752
           ||||||||||||||||  || |||||||||   |||||||||
Db   1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy   753 TTCGGAGCGGCTGAAGGAGGCTCATGTGAAGGCTTCAGTGTCACAGTCCGCTGCACATCACA 812
           ||||||||| ||||||   |  |||||   |||  |
Db   1064 TCCGGAGCGCTGAAGATGACTGTTGTGAAGCTGAATTTCCCACGGGAGTGTCAGCAGCC 1123

Qy   813 ACATTACTTAACAGAACAGTCACCGACACAACATGCTGTGCTGGAGACACTCGGAGCGG 872
           |||||||||  |||||||||||||||||  |||||||||||||
Db   1124 CCACTACTACGGCTCTGAAGTCACCACCAAAAATGCTGTGTGCTGACCCACAGTG--- 1180

Qy   873 CGGGCCCCAGGCAAACTTGCACGACGGCGTGCCAGGGCGAATTCCGGAGGCCCCCTGGTGTG 932
           |          |||||||||||||||||||||||||||||||||||
Db   1181 --------------GAAAACAGATTCCTGCCAGGGAGGACTCAGGGGGACCCCTCGTCTG 1225

Qy   933 TCTGAACGATGGCCGCATGACTTGGTTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACA 992
           ||||||||||| |||||
Db   1226 TTCCCTCCAAGGCCGCATGACTTGACTGGAATTGTGAGCTGGGCCGTGGGATGTGCCCT 1285

Qy   993 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATTCG 1046
           ||||||    ||
Db   1286 GAAGGACAAGCCAGGCGTCTCACGAGTCTCACACTTCTTACCCTGGATCCG 1339
```

Search completed: August 19, 2003, 14:22:30
Job time : 302.529 secs

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 22:44:33 ; Search time 3115.64 Seconds
                                          (without alignments)
                                          8307.845 Million cell updates/sec

Title: US-09-987-455-7
Perfect score: 1065
Sequence: 1 tctgaggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215223856 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
    1: em_estba:*
    2: em_esthum:*
    3: em_estin:*
    4: em_estmu:*
    5: em_estov:*
    6: em_estpl:*
    7: em_estro:*
    8: em_htc:*
    9: gb_est1:*
   10: gb_est2:*
   11: gb_htc:*
   12: gb_est3:*
   13: gb_est4:*
   14: gb_est5:*
   15: em_estfun:*
   16: em_estom:*
   17: em_gss_hum:*
   18: em_gss_inv:*
   19: em_gss_pln:*
   20: em_gss_vrt:*
   21: em_gss_fun:*
   22: em_gss_mam:*
   23: em_gss_mus:*
   24: em_gss_pro:*
   25: em_gss_rod:*
   26: em_gss_phg:*
   27: em_gss_vrl:*
   28: gb_gss1:*
   29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%Query Match	Length	DB	ID	Description	
c 1	895	84.0	930	13	BX329047	BX329047	BX329047
2	853	80.1	912	13	BU149958	BU149958	AGENCOURT
3	850.6	79.9	962	13	BX389609	BX389609	BX389609
4	818.6	76.9	891	13	BU146167	BU146167	AGENCOURT
5	815.8	76.6	993	13	BQ278177	BQ278177	AGENCOURT
6	811.2	76.2	883	13	BQ690978	BQ690978	AGENCOURT
7	807.6	75.8	918	13	BX389608	BX389608	AGENCOURT
8	796.2	74.8	930	13	BU845014	BU845014	AGENCOURT
9	782.8	73.5	888	13	BU157720	BU157720	AGENCOURT
10	767.4	72.1	916	12	BI765247	BI765247	603050338
11	765.2	71.8	950	13	BQ927899	BQ927899	AGENCOURT
12	706.8	66.4	948	10	BE616613	BE616613	601278864
13	697.4	65.5	958	12	BG763582	BG763582	602735987
14	694.4	65.2	974	13	BQ960216	BQ960216	AGENCOURT
15	693.4	65.1	709	10	BE732704	BE732704	601571366
16	685.2	64.3	1060	12	BM556999	BM556999	AGENCOURT
17	677	63.6	1042	10	BE732413	BE732413	601569644
18	672	63.1	957	13	BQ687538	BQ687538	AGENCOURT
19	668.8	62.8	948	13	BQ687779	BQ687779	AGENCOURT
20	665.2	62.5	704	9	AU134301	AU134301	AU134301
21	658.2	61.6	903	13	BQ689095	BQ689095	AGENCOURT
22	656.2	61.6	830	13	BQ896938	BQ896938	AGENCOURT
23	655.2	61.5	904	13	BQ689840	BQ689840	AGENCOURT
24	655.2	61.5	911	13	BQ684734	BQ684734	AGENCOURT
25	654.6	61.5	867	13	BU179903	BU179903	AGENCOURT
26	652.4	61.3	818	9	AU124602	AU124602	AU124602
27	651	61.1	818	12	BM006474	BM006474	603615409
28	649.6	61.0	797	13	BQ879911	BQ879911	AGENCOURT
29	642.2	60.3	1037	13	BU157493	BU157493	AGENCOURT
30	631.2	59.3	813	10	BG120780	BG120780	602348540
31	629.8	59.1	663	9	AU134463	AU134463	AU134463
32	624	58.6	624	10	BE549489	BE549489	601278502
33	585.4	55.0	587	9	BE394944	BE394944	601308225
34	561	52.7	561	9	AL047284	AL047284	DKFZp586H
35	553.4	52.0	555	14	CB128855	CB128855	K-EST0178
36	535.4	50.3	537	14	CB155052	CB155052	K-EST0213
37	531	49.9	964	12	BG769384	BG769384	602742678
38	517.4	48.6	905	13	BU859271	BU859271	AGENCOURT
39	514.8	48.3	883	14	CD359959	CD359959	AGENCOURT
40	514.4	48.3	516	14	CB140923	CB140923	K-EST0194
41	496.4	46.6	916	13	BU157241	BU157241	AGENCOURT
42	487.8	45.8	527	9	AW630119	AW630119	hh78e06.y
43	485.8	45.6	553	9	AW394127	AW394127	MR2-TT001
44	477.8	44.9	917	13	BU190904	BU190904	AGENCOURT
45	475.8	44.7	937	13	BQ679501	BQ679501	AGENCOURT

ALIGNMENTS

RESULT 1
BX329047/c
LOCUS   BX329047   930 bp   mRNA   linear   EST 01-MAY-2003
DEFINITION BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 3-PRIME, mRNA sequence.
ACCESSION   BX329047
VERSION   BX329047.1 GI:30307809
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4498.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAIO212D08_C
            S01980.1&cluster=4498.r. Contact : Feng Liang Email :
            fliang@lifetech.com URL : http://fulllength.invitrogen.com/
            InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID :

```
FEATURES
     source          CS0BAI021ZD08 CS01980 1.
                     Location/Qualifiers
                     1..930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DIO39YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT       184 a       268 c       282 g       194 t       2 others
ORIGIN

Query Match            84.0%;     Score 895;    DB 13;    Length 930;
Best Local Similarity  98.6%;     Pred. No. 5.7e-223;
Matches  912;  Conservative   0;  Mismatches  12;  Indels    1;  Gaps     1;

Qy       1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCAGCGACACGCACAGC       60
           |||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     927 TTTGAGGGAAACAGTGACTGCTAC-TTGGGAATGGGTCAGCTCAGCGACGACACGGGACAGC   869

Qy      61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT      120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     868 CTCACCGAGTCGGGTGCTNCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   809

Qy     121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCACATAATTACTGCGGG      180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     808 TACACAGCACAGAACNCCAGTGCCCAGGCACTGGGCCTGGGCACATAATTACTGCGGG   749

Qy     181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG      240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     748 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   689

Qy     241 GAGTACTGTGATGTGCCCCTCCTGTCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG      300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     688 GAGTACTGTGATGTGCCCCTCCTGTCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   629

Qy     301 TTTCGGCATCAAAGGAGGGCTCTTCGCCGCACATCGCCTTCGCCGACATCGCCCCTGCCATC      360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 TTTCGGCATCAAAGGAGGGCTCTTCGCCGCACATCGCCTTCGCCGACATCGCCCCTGCCATC   569

Qy     361 TTTGCCAAGCACGAGGAGGTGCCCGGAGAGGTTCGTGCGGGGCATACTCATCAGC      420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     568 TTTGCCAAGCACCAGGAGGTGCCCGGAGAGGTCGTGCGGGGCATACTCATCAGC   509

Qy     421 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCTG      480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCTG   449

Qy     481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA      540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA   389

Qy     541 GTCGAAAAAATAACATTGCTCCATAAGGAATTCGATGATGATGACACTTACGACACATTGCGG      600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 GTCGAAAAAATAACATTGCTCCATAAGGAATTCGATGATGATGACACTTACGACACATTGCGG   329

Qy     601 CTTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCGCGTGCCCAGGAGCAGCGTGTCCGACT      660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 CTTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCGCGTGCCCAGGAGCAGCGTGTCCGACT   269

Qy     661 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACGGAGTTGAGCTTCCGGGC      720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACGGAGTTGAGCTTCTCCGGGC   209

Qy     721 TACGGCAAGCATGAGGCCTTGTCTTCCTTGTCTCATTCGGAGCGGCTGAAGGAGGCTCATGTC      780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 TACGGCAAGCATGAGGCCTTGTCTTCCTTGTCTCATTCGGAGCGGCTGAAGGAGGCTCATGTC   149

Qy     781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC      840
```

```
Db     148 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC       89
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGGGCGGGGGCCCCCAGGCAAACTTGCACGACGCC      900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGGGCGGGGGCCCCCAGGCAAACTTGCACGACGCC       29

Qy     901 TGCCAGGGCGATTCGGGAGGCCCCC      925
           |||||||||||||||||||||||||
Db      28 TGCCAGGGCGATTCGGGAGGGCCCCC        4
```

```
RESULT 2
BU149958
LOCUS          BU149958               912 bp   mRNA      linear     EST 03-SEP-2002
DEFINITION     AGENCOURT 8137213 Lupski_dorsal_root_ganglion Homo sapiens cDNA
               clone IMAGE:6184119 5', mRNA sequence.
ACCESSION      BU149958
VERSION        BU149958.1  GI:22663490
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               1 (bases 1 to 912)
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
  AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE        Unpublished
  JOURNAL      Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Dr. James R. Lupski
COMMENT        cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Agencourt Bioscience Corporation
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLAM13572  row: m  column: 16
               High quality sequence start: 18
               High quality sequence stop: 644.
FEATURES       Location/Qualifiers
     source    1..912
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6184119"
               /sex="male"
               /tissue_type="dorsal root ganglia"
               /dev_stage="adult, 36 yr"
               /lab_host="DH10B"
               /clone_lib="Lupski dorsal root ganglion"
               /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
               NotI; Site 2: SalI; cDNA made by oligo-dT priming.
               Directionally cloned using the following adaptors:
               5'-TCGACCCACGCGTCCG-3' and
               5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
               1 kb for average insert length 1.7 kb. This is a primary
               library, non-amplified. Library constructed by Life
               Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
               College of Medicine) and is available through Life
               Technologies."
BASE COUNT      189 a      272 c      271 g      179 t      1 others
ORIGIN

Query Match            80.1%;     Score 853;    DB 13;    Length 912;
Best Local Similarity  98.9%;     Pred. No. 5.4e-212;
Matches  901;  Conservative   0;  Mismatches   5;  Indels    5;  Gaps     4;

Qy     131 AGAACCCCAGTGCCCAGGCACTGGG--CCTGGGCAAACATCATAATTACTGCGGAATCCTGA      188
           ||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db       1 AGAACCCCAGTGCCCAGGCACTGGGCCCTGGGCAAACATCATAATTACTGCGGAATCCTGA       60

Qy     189 TGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTG      248
```

AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen. This sequence belongs to sequence cluster
           4498.r For more information about this cluster, see
           http://www.genoscope.cns.fr/
           cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_3&cluster=4498.r.
           Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_3.

FEATURES   Location/Qualifiers
     source   1..962
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI03YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT   203 a   290 c   277 g   191 t   1 others
ORIGIN

Query Match   79.9%; Score 850.6; DB 13; Length 962;
Best Local Similarity 96.4%; Pred. No. 2.3e-211;
Matches   891; Conservative 0; Mismatches 30; Indels 3; Gaps 2;

Qy   142 GCCAGGCACTGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGCCAAG 201
         |||||||||||| ||||| ||||| |||||||||||||||||||||||||||||||||
Db   1 GCCCAGGCACTGGNGGCCTGGGCAAACTAATTACTGCCGGAATCCTGATGGGGATGCCAAG 60

Qy   202 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGGGAGTACTGGGAGTGTGATGTGCCCTCC 261
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGGGAGTACTGGGAGTGTGATGTGCCCTCC 120

Qy   262 TGCTCCACCTGCGCGCCTGAGACAGTAGCAGCCAGCCTCAGTTTCGCATCAAAGAGAGGCTC 321
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTCCACCTGCGCGCCTGAGACAGTAGCAGCCAGCCTCAGTTTCGCATCAAAGAGAGGGTC 180

Qy   322 TTCGCCGACATCGCCTCCCACCCCTGGCGAGGCTGCCATCTTTGCCAAGCACAGGAGGTCG 381
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTCGCCGACATCGCCTCCCACCCCTGGCGAGGCTGCCATCTTTGCCAAGCACAGGAGGTCG 240

Qy   382 CCCGGAGAGGCGGTTCCGAGGAGGAGGAGCATACTACAGTTTCGGCTGATTCTCTCTGCC 441
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCCGGAGAGGCGGTTCCGAGGAGGAGGAGCATACTACAGTTTCGGCTGATTCTCTCTGCC 300

Qy   442 GCCCACTGCTTCCAGGAGGTTTCCGCCCCCACCCACCACCTGACGGTGATCTTGGGCAGAACA 501
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCCCACTGCTTCCAGGAGGTTTCCGCCCCCACCCACCACCTGACGGTGATCTTGGGCAGAACA 360

Qy   502 TACCGGGTGGTCCCTGGCGAGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCAT 561
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TACCGGGTGGTCCCTGGCGAGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCAT 420

Qy   562 AAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGAT 621
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGAT 480

Qy   622 TCGTCCCGTGTCCCAGAGACAGCCGTGGTCGTGCCTTCCCCCGGCGGAC 681
         ||||||||||||||||||||| ||||||||||||||||||||||||||
Db   481 TCGTCCCGTGTCCCAGAGAGCAGCCGTGGTCGTGCCTTCCCCCGGCGGAC 540

Qy   682 CTGCAGTGCCGACTGCGACGAGTGTGAGCTCTCCGGCTACCGCAAGCATGAGGGCCTTG 741
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGCAGTGCCGACTGCGACGAGTGTGAGCTCTCCGGCTACCGCAAGCATGAGGGCCTTG 600

Db   61 TGGGGATGCCCAGGCACCCTGCGTGGTGCCACCTGCGCCGCTGAGAAGAACCGCAGCGTGGGAGTACTG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   249 TGATGTGCCCTCCTGCTCCACCTGCGCCGCGGCGGCTCGAGCAGTTTCGCAT 308
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGATGTGCCCTCCTGCGTCCACCTGCGGCGGCTCGAGCAGTTTCGCAT 180

Qy   309 CAAAGGAGGGCTCTTCGCCGACATCGCCGCTGCCTCCACCCCTGGCGAGGCTGCCATCTTTGCCAA 368
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAAAGGAGGGCTCTTCGCCGACATCGCCGCTGCCTCCACCCCTGGCGAGGCTGCCATCTTTGCCAA 240

Qy   369 GCACAGGAGGTCGCCCGGAGAGGCGGTTCTGTGCGGGGGCATATCATCAGCTCCTGCTG 428
         ||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db   241 GCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCTGCTG 300

Qy   429 GATTCTCTCGCCGCCGCCACTGCTTCCAGGAGAGGTTCCGCCCCCACCACCTGACGTGAT 488
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GATTCTCTCGCCGCCCACTGCTTCCAGGAGAGGTTCCGCCCCCACCACCTGACGGTGAT 360

Qy   489 CTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAGAAATTTGAAGTCGAAAA 548
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAGAAATTTGAAGTCGAAAA 420

Qy   549 ATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCA 608
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGGCTGCTGCA 480

Qy   609 GCTGAAATCGATTCGTCCCCGTGTCCCGCGTGTCCCAGAGAGCAGCGTGGTCGTGCCTGTGTGCCT 668
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCTGAAATCGATTCGTCCCCGTGTCCCGCGTGTCCCAGAGAGCAGCGTGGTCGTGCCTGTGTGCCT 540

Qy   669 TCCCCCGGCGGACCTGCAGCTGCCGGACTGGAGCTGCGACGAGTGTGAGCTCTCCGGCTACGGCCAA 728
         ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   541 TCCCCCGGCGGACCTGCAGCTGCCGGACTGCGACGAGTGTGAGCTCTCCGGCTACGGCCAA 600

Qy   729 GCATGAGGGCCTTGTCTCCTTCTCTTCTATTCGGAGGAGGCTGAAGGAGGCTCATGTCAGACTGTA 788
         |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   601 GCATGAGGGCCTTGTCTCCTTCTCCTTTCTATTCGGAGGAGGCTGAAGGAGGCTCATGTCAGACTGTA 660

Qy   789 CCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCT 848
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCT 720

Qy   849 GTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCAGACGCGCCTGCCAGGG 908
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCAGACGCGCCTGCCAGGG 780

Qy   909 -CGATTCGGGGAGG-CCCCCGTGGGCTGTGTGGCTGTGTCTGAACGATGGCCGCATGACTTTGTGGGGCATC 966
          ||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db   781 NCGATTCGGGGAGGCGTTCGGGGAGGCGTTCGGGGTGTCCGATGGCCGCATGACTTTGTGGGGCATC 840

Qy   967 ATCAGCTGGGGGCCTGGGCGTGTGGACAAGGATGT-CCCGGGTGTGTACACCAAGGTTAC 1025
         ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   841 ATCAGCTGGGGGCCTGGGCGTGTGGGACAAGGATGTGCCCGGGTGTGTACACCAAGGTTAC 900

Qy 1026 CAACTACCTAG 1036
         | ||||||||
Db   901 AAACTACCTAG 911

```
BASE COUNT      178 a    266 c    271 g    176 t
ORIGIN                Technologies."

Query Match          76.9%;   Score 818.6;  DB 13;   Length 891;
Best Local Similarity 98.2%;   Pred. No. 5.2e-203;
Matches  849;  Conservative   0;  Mismatches  14;  Indels   2;  Gaps   2;

Qy    136  CCCAGTGCCCAGGCACTGGGCTCGGGCAAACATAATTACTGCCGGAATCCTGATGGGGAT  195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  CCCAGTGCCCAGGCACTGGGCTCGGGCAAACATAATTACTGCCGGAATCCTGATGGGGAT   60

Qy    196  GCCAAGCCTGGTGCCACGTGCTGAAGAACCGGAGGCTGACGTGGAGTACTGTGATGTG  255
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GCCAAGCCTGGTGCCACGTGCTGAAGAACCGGAGGCTGACGTGGAGTACTGTGATGTG  120

Qy    256  CCCTCTGCTCACCTGCGGCCTGAGACAGCAGCCTCAGTTTCGCATCAAAGGA  315
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CCCTCCTGCTCCACCTGCGGCCTGAGACAGTCAGCAGCCTCAGTTTCGCATCAAAGGA  180

Qy    316  GGGCTCTTCGCCGACATCGCCTTCGCCAGGTGCCATCTTGCCAAGCACAGG  375
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  GGGCTCTTCGCCGACATCGCCTTCCCACCCTGCAGGCTGCCATCTTTGCCAAGCACAGG  240

Qy    376  AGTCGCCGGAGAGCGGGTTCCTGTGCGGGGCATACTCATCAGCTCATCAGCTCGATTCTC  435
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  AGTCGCCGGAGAGCGGGTTCCTGTGCGGGGCATACTCATCAGCTCATCAGCTCGATTCTC  300

Qy    436  TCTGCCGCCACTGCTTCCAGGAGAGTTTCGGCCCACCACCTGACGGTGATCTTGGGC  495
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TCTGCCGCCACTGCTTCCAGGAGAGTTTCCCCACCACCTGACGGTGATCTTGGGC  360

Qy    496  AGAACATACCGGTGGTCCTGGCGAGGAGGCAGAAATTTGAAGTCGAAAATACATT  555
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGAACATACCGGTGGTCCTGGCGAGGAGGCAGAAATTTGAAGTCGAAAATACATT  420

Qy    556  GTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAA  615
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAA  480

Qy    616  TCGGATTCGTCCCGCTGGTCCCGCTGGCCTGTGGCCTTCCCCCG  675
                |||||||||||||||||||||||||||||||||||||||||||
Db    481  TCGGATTCGTCCCGCTGGTCCCGCTGGCCTGTGGCCTTCCCCCG  540

Qy    676  GCGGACCTGCAGCTGCCGGACTGGACGGCGGCTCCGGCTCACGGCCAAGCATGAG  735
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GCGGACCTGCAGCTGCCGGACTGGACGGCGGCTCTCCGGCTCACGGCCAAGCATGAG  600

Qy    736  GCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCC  795
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCC  660

Qy    796  AGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCT  855
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCT  720

Qy    856  GGAGACACTCGGAGCGGC-GGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGCGGATTC  914
                ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    721  GGAGACACTCGGAGCGGCCGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGCGGATTC  780

Qy    915  GGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCAT-GACTTTGGTGGGCATCATCAGCT  973
                |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    781  GGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGGACTTTGGTGGGCATCATCAGCT  840

Qy    974  GGGGCCTGGGCTGTGTGACAGAGGA  998
                |||||||||||||||||||||||||
Db    841  GGGGGCCCCTGGGCTGGGGGGACAGA  865

Qy    742  TCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGC  801
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGC  660

Qy    802  TGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGAC  861
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  TGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGAC  720

Qy    862  ACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGCGGATTCGGGAGGC  921
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGCGGATTCGGGAGGC  780

Qy    922  CCCCTGGTGTGTCTGAACGATGCCGCGATGACTTTGGTGGGCATCATCAGCTGGGGCCTG  981
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  CCCCTGGTGTGTCTGAACGATGCCGCGATGACTTTGGTGGGCATCATCAGCTGGGGCCTG  840

Qy    982  GGCTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGG  1041
                |||||||| ||| |||||||| ||||         ||| ||| ||||
Db    841  GGCTGTGTGGAACAGAAGATGTCCCGGGTTGTTCCCACCAAG--TTACCACTTACTAAACTGG  898

Qy   1042  ATTCGTGACAACATGCGACCGTGA  1065
                |||| |||  ||||||||| ||
Db    899  ATTCGGGAC-ACATGCGACCGGGA  921
```

RESULT 4
BU146167
LOCUS   BU146167   891 bp   mRNA   linear   EST 03-SEP-2002
DEFINITION   AGENCOURT_7981262 Lupski_dorsal_root_ganglion Homo sapiens cDNA
             clone IMAGE:6185661 5', mRNA sequence.
ACCESSION   BU146167
VERSION   BU146167.1   GI:22659699
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 891)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM13576 row: m column: 22
             High quality sequence stop: 665.
FEATURES   Location/Qualifiers
     source   1..891
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6185661"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGGACGCGCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life

RESULT 5
BQ278177
LOCUS   BQ278177   993 bp   mRNA   linear   EST 07-MAY-2002
DEFINITION   AGENCOURT_7061709 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5805153
```

```
ACCESSION   BQ278177
VERSION     BQ278177.1  GI:20488385
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 993)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2043  row: o  column: 10
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1..993
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5805153"
                     /tissue_type="teratocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
                     /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH MGC Library."
BASE COUNT      235 a    295 c    269 g    194 t
ORIGIN
Query Match         76.6%;  Score 815.8;  DB 13;  Length 993;
Best Local Similarity  99.6%;  Pred. No. 2.9e-202;
Matches  828;  Conservative  0;  Mismatches  2;  Indels  1;  Gaps  1;

Qy    234 GACGTGGGAGTACTGTGATGTGCCCTCCCTGTCCACCTGCGGCTCGAGACAGTACAGCCA 293
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      1 GACGTGGGAGTACTGTGATGTGCCCTCCCTGCT-CACCTGCGGCTCGAGACAGTACAGCCA 59

Qy    294 GCCTCAGTTCGGCATCAAAGGAGGGCTCTTCGGCCGACATCGCCTCCCACCCTGGCAGGC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GCCTCAGTTCGGCATCAAAGGAGGGCTCTTCGGCCGACATCGCCTCCCACCCTGGCAGGC 119

Qy    354 TGCCATCTTCGCCAAGCACAGAGAGGTCGCCCGAGAGCGGGTTCCTGTGCGGGGGCATACT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 TGCCATCTTCGCCAAGCACAGAGAGGTCGCCCGAGAGCGGGTTCCTGTGCGGGGGCATACT 179

Qy    414 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCCA 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCCA 239

Qy    474 CCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAA 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAA 299

Qy    534 ATTTGAAGTCGAAAAATACATTGTCCATAAGAATTCGATGATGACACTTACGACACAATGA 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 ATTTGAAGTCGAAAAATACATTGTCCATAAGAATTCGATGATGACACTTACGACACAATGA 359

Qy    594 CATTGGCCTGCTGCAGCTGAAATGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGT 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                                                                   Db    360 CATTGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGT 419
                                                                   Qy    654 CCGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCGGACCTGCGGACCTGAGGCGGCTGAGGAGGC 713
                                                                             ||||||||||| |||||||||||||||||||||||||||||| ||||||||
                                                                   Db    420 CCGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCGGACCTGCGGACCTGGACGGAGGTGAGCT 479
                                                                   Qy    714 CTCCGGCTACGGCAAGACATGAGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGGAGGC 773
                                                                             |||||||||||||||||||||||||||||||||||||||| ||||||||
                                                                   Db    480 CTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGGAGGC 539
                                                                   Qy    774 TCATGTCAGACTGTACCCATCCAGCCGCTGACATCACAACATTACTTAACAGAACAGT 833
                                                                             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||
                                                                   Db    540 TCATGTCAGACTGTACCCATCCAGCCGCTGACATCACAACATTACTTAACAGAACAGT 599
                                                                   Qy    834 CACCGACAACATGTCGTGTCTGAGACACTCGGAGCGCGGCGGGCCCAGGCAACTTGCA 893
                                                                             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                   Db    600 CACCGACAACATGTCGTGTCTGAGACACTCGGAGCGCGGCGGGCCCAGGCAAACTTGCA 659
                                                                   Qy    894 CGACGCCTGCCAGGGGATTCGGGAGGCCCCGTGGTGTGTCTGAACGATGGCCGCCATGAC 953
                                                                             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                   Db    660 CGACGCCTGCCAGGGGATTCGGGAGGCCCCGTGGTGTGTCTGAACGATGGCCGCCATGAC 719
                                                                   Qy    954 TTTGGTGGGCATCATCAGCTGGGGCCTGGCTGTGGACAGAAGGATGTCCCGGGGTGTA 1013
                                                                             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                   Db    720 TTTGGTGGGCATCATCAGCTGGGGCCTGGCTGTGGACAGAAGGATGTCCCGGGGTGTA 779
                                                                   Qy   1014 CACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCGTG 1064
                                                                             |||||||||||||||||||||||||||||||||||||||||||||||||
                                                                   Db    780 CACCAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCGTG 830
```

```
RESULT 6
BQ690978
LOCUS       BQ690978  AGENCOURT_8046707  NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
                      5', mRNA sequence.                              883 bp    mRNA    linear    EST 15-JUL-2002
DEFINITION  AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
            5', mRNA sequence.
ACCESSION   BQ690978
VERSION     BQ690978.1  GI:21816294
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2367  row: o  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1..883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6209120"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
                     Site 2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
```

Right column

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
            Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 918)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4498.r For more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_2&cluster=4498.r.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_2.
FEATURES    Location/Qualifiers
source      1..918
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DIO39YK15"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with NotI and cloned into the NotI and EcoRI V
            sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT  189 a   273 c   264 g   186 t   6 others
ORIGIN

Query Match          75.8%;   Score 807.6;  DB 13;  Length 918;
Best Local Similarity 96.4%;  Pred. No. 4e-200;
Matches 855;  Conservative 0;  Mismatches 29;  Indels 3;  Gaps 3;

Qy  170 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCTGCCACGTGCTGAAGAACCGCA 229
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCTGCCACGTGCTGAAGAACCGCA 76

Qy  230 GGCTGACGTGGGAGTACTGTGATGTGCCGGCCTGAGACAGTACA 289
         ||||||||||||||||||||||||||||||||||||||||||||
Db   77 GGCTGACGTGGGAGTACTGTGATGTGCCGGCCTGAGACAGTACA 136

Qy  290 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 349
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 196

Qy  350 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGGCCCGGAGAGAGCCGGTTCCTGTGCGGGGGCA 409
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGGCCCGGAGAGAGCCGGTTCCTGTGCGGGGGCA 256

Qy  410 TACTCATCAGCTCCTGGATTCTCTGCGCCCACTGCTTCTGCCAGGAGGTTTCCGCCC 469
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 TACTCATCAGCTCCTGGATTCTCTGCGCCCACTGCTTCTGCCAGGAGGTTTCCGCC 316

Qy  470 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGC 529
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGC 376

Qy  530 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 589
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  377 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 436

Qy  590 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCTGCCCAGGAGAGCAGCG 649
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  437 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCTGCCCAGGAGAGCAGCG 496

Qy  650 TGGTCCGGACCGTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGGACGGCGAGTGTG 709
```

Left column

```
            (Stratagene) and Superscript II RT (Life Technologies)."
            Note: this is a NIH_MGC Library.
BASE COUNT  184 a   265 c   259 g   175 t
ORIGIN

Query Match          76.2%;   Score 811.2;  DB 13;  Length 883;
Best Local Similarity 99.0%;  Pred. No. 4.5e-201;
Matches 816;  Conservative 0;  Mismatches 0;  Indels 8;  Gaps 0;

Qy  112 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAAT 171
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAAT 60

Qy  172 TACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 231
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy  232 CTGACGTGGGAGTACTGTGATGTGCCGGCCTGGGGCCTGAGACAGTACAGC 291
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGACGTGGGAGTACTGTGATGTGCCCTCCAACCTGCGGCCTGAGACAGTACAGC 180

Qy  292 CAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAG 351
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAG 240

Qy  352 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGAGCCGGTTCCTGTGCGGGGGCATA 411
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGAGCCGGTTCCTGTGCGGGGGCATA 300

Qy  412 CTCATCAGCTCCTGGATTCTCTGCGCCCACTGCTTCTGCCAGGAGGTTTCCGCCCC 471
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CTCATCAGCTCCTGGATTCTCTGCGCCCACTGCTTCTGCCAGGAGGTTTCCGCCC 360

Qy  472 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGCAG 531
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGCAG 420

Qy  532 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 591
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 480

Qy  592 GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCAGCGTG 651
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCAGCGTG 540

Qy  652 GTCCGCACTGTGTGCCTTCCCCCGGCCGGACCTGCAGCTGGACGCGGAGTGTGAG 711
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GTCCGCACTGTGTGCCTTCCCCCGGCCGGACCTGCAGCTGGACGCGGAGTGTGAG 600

Qy  712 CTCTCCGGCTACGGCAAGCATGAGGCCGTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAG 771
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAG 660

Qy  772 GCTCATGTCAGACTGTACCAGCCGCTGCACATCACAAACATTTACTTAACAGAACA 831
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GCTCATGTCAGACTGTACCAGCCGCTGCACATCACAAACATTTACTTAACAGAACA 720

Qy  832 GTCACCGACAACATGTCGTGTCGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTG 891
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GTCACCGACAACATGTCGTGTCGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTG 780

Qy  892 CACGACGCCTGCCAGGCGGATTCGGGAGGGCCCCCCTGTGTGTCT 935
         |||||||||||||||||||||||||||||||||| ||
Db  781 CACGACGCGCCTGCCAGGCGGAATTCGGGAGGGCCCCCCTGTGTGTGT 824

RESULT 7
BX389608
LOCUS       BX389608   918 bp   mRNA   linear   EST 08-MAY-2003
DEFINITION  BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DIO39YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389608
VERSION     BX389608.1   GI:30463469
```

ORIGIN

```
Query Match        74.8%;     Score 796.2;  DB 13;    Length 930;
Best Local Similarity   99.4%;     Pred. No. 3.8e-197;
Matches   820; Conservative    0; Mismatches    3;  Indels    2;  Gaps    2;

Qy    234 GAGCTGGGAGTACTGTGATGTGCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCA 293
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GACGTGGGAGTACTGTGATGTGCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCA 60

Qy    294 GCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCAGGC 353
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCAGGC 120

Qy    354 TGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGGCGGTTCCTGTCGTGGGGGCATACT 413
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGGCGGTTCCTGTCGTGGGGGCATACT 180

Qy    414 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 473
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 240

Qy    474 CCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGCAGAGA 533
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGCAGAGA 300

Qy    534 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 593
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 360

Qy    594 CATTGCCGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCAGGAGAGCAGCGTCGT 653
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATTGCCGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCAGGAGAGCAGCGTGGT 420

Qy    654 CCGCACTGTGTGCCTTCCCCCCGCGGGACCTGCAGCTGGACTGGACTGGAGTGTGAGCT 713
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCGCACTGTGTGCCTTCCCCCCGCGGGACCTGCAGCTGGACTGGACTGGAGTGTGAGCT 480

Qy    714 CTCCGGCTACGGCAAGCAATGAGGGCTTGTCTCCTTCCTTTCTATTCGGAGCGCTGAAGGAGGC 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCCGGCTACGGCAAGCAATGAGGGCTTGTCTCCTTCCTTTCTATTCGGAGCGCTGAAGGAGGC 540

Qy    774 TCATGTCAGACTGTACCCATCAGCGCGTGCACATCAACATTTACTTAACAGAACAGT 833
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCATGTCAGACTGTACCCATCAGCGCGTGCACATCAACATTTACTTAACAGAACAGT 600

Qy    834 CACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCAGGCAAACTTGCA 893
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCAGGCAAACTTGCA 660

Qy    894 CGACGCCTGCCAGGGCGATTCGGGGAGGCCCCCTGGTGTGTGTCTGAACGATGGCCGCATGAC 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGACGCCTGCCAGGGCGATTCGGGGAGGCCCCCTGGTGTGTGTCTGAACGATGGCCGCATGAC 720

Qy    954 TTTGGTGGGCATCATCAGCTCGGGGCCTGGGCGTGTGGACAGAGGATGTCCC-GGGTGTGT 1012
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db    721 TTTGGTGGGCATCATCAGCTCGGGGCCTGGGCGTGTGGACAGAGGATGTCCCGGGGTGTGT 780

Qy   1013 ACA-CAAAGGTTACCAACTACCTTAGACTGGATTCGTGACAACATG 1056
              ||| | ||||||||||||||||||||||||||||||||||||||
Db    781 ACACCCAAGGGTACCAACTACCTTAGACTGGATTCGTGACAACATG 825
```

RESULT 9
BU157720

LOCUS   BU157720
DEFINITION AGENCOURT_8043192 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084476
            5', mRNA sequence.
ACCESSION   BU157720
VERSION   BU157720.1 GI:22671252
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens

```
Db    497 TGGTCCGCACTGTGTGCCTTCCCCCCGCGGGACCTGCAGCTGGACTGGACTGGAGTGTGTG 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    710 AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTCCTTTCTATTCGGAGCGGCTGAAGG 769
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTCCTTTCTATTCGGAGCGGCTGAAGG 616

Qy    770 AGGCTCATGTCAGACTGTACCCATCAGCCGCGTGCACATCAACAACATTTACTTAACAGAA 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    617 AGGCTCATGTCAGACTGTACCCATCAGCCGCGTGCACATCAACAACATTTACTTAACANGA 676

Qy    830 CAGTCACCGACAACACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACT 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    677 CAGTCACCGANCACATGCTGTGTGCTGGAGACACTCGNAGCGCGGCGGNCCCCA-GCAAACT 735

Qy    890 TGCACGACGCCTGCCAGGGCGATTCGGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 949
              |||||||||||||||||||||||||||||||||||||||||| |||||||
Db    736 TGCACGACGCCCTGCCAGGGCGATTNCGGAGGGCCCCCTGGTGTGTCTGACGATTGCCCAA 795

Qy    950 TGA-CTTTGGTGGGCATCATCAGCTGGGG--CCTGGGCGTGTGGACAGAGAGGATGTCCGGG 1007
              ||| ||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db    796 TGACCTTTGGTGGCCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAGAGGATGTTCCCGG 855

Qy   1008 TGTGTACACAAAGGTTACCAACTACCTTAGACTGGATTCGTGACAACA 1054
              ||| ||| |||||| ||||||||||||||||||||||||||
Db    856 TGTGTCACCCAAGTTTACAACTACCTTAAACCGGATTCGGGACAACA 902
```

RESULT 8
BU845014

LOCUS   BU845014   930 bp   mRNA   linear   EST 16-OCT-2002
DEFINITION AGENCOURT 10412263 NIH_MGC_109 Homo sapiens cDNA clone
            IMAGE:6578533 5', mRNA sequence.
ACCESSION   BU845014
VERSION   BU845014.1 GI:24029455
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2780 row: g column: 13
            High quality sequence stop: 639.
FEATURES   Location/Qualifiers
     source   1..930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6578533"
                     /tissue_type="teratocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT   199 a   279 c   266 g   186 t
```

```
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM2316 row: e  column: 21
             High quality sequence stop: 630.
FEATURES     Location/Qualifiers
source       1..888
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6084476"
             /tissue_type="ductal carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH MGC 110"
             /note="Organ: pancreas; Vector: pQTB7; Site_1: XhoI;
             Site_2: EcoRI; cDNA made by oligo-dT priming.
             Directionally cloned into EcoRI/XhoI sites using the
             following 5' adaptor: GGCACGAG(G). Library constructed by
             Ling Hong in the laboratory of Gerald M. Rubin (University
             of California, Berkeley) using ZAP-cDNA synthesis kit
             (Stratagene) and Superscript II RT (Life Technologies).
             Note: this is a NIH MGC Library."
BASE COUNT      181 a     265 c     270 g     172 t
ORIGIN
```

```
Query Match       73.5%;  Score 782.8;  DB 13;  Length 888;
Best Local Similarity   97.2%;  Pred. No. 1.2e-193;
Matches   861;  Conservative    0;  Mismatches   17;  Indels    8;  Gaps   6;

Qy    112 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTCGGCCTGGCAAACATAAT 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTCGGCCTGGCAAACATAAT 60

Qy    172 TACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy    232 CTGACGTGGGAGTACTGTGATGTGCCTCCTGCTCCACCTGCGGCGTCCACCTGAGACAGC 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGACGTGGGAGTACTGTGATGTGCCTCCTGCTCCACCTGCGGCGTCCACCTGAGACAGC 180

Qy    292 CAGCCTCAGTTTCGACATCAAAGAGGAGGGCTCTTCGCCGCCACATCGCCTCCACCCCTGCAG 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGCCTCAGTTTCGACATCAAAGAGGAGGGCTCTTCGCCGCCACATCGCCTCCACCCCTGCAG 240

Qy    352 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATA 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATA 300

Qy    412 CTCATCAGCTCCTGCTGATTCTCTGCCGCCCACTGCTTCCAGGAGAGAGGTTTCCGCCC 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTCATCAGCTCCTGCTGATTCTCTGCCGCCCACTGCTTCCAGGAGAGAGGTTTCCGCCC 360

Qy    472 CACCACCTGACCGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCCTGGAGGAGGAGCAG 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACCTGACCGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCCTGGAGGAGGAGCAG 420

Qy    532 AAATTTGAAGTCCAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 480
```

```
Qy    592 GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCACTGTGGGAGGCAGCGTG 651
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCACTGTGGGAGGCAGCGTG 540

Qy    652 GTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGACTGCAGCTGCCGACGAGTGTGAG 711
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGACTGCAGCTGCCGACGAGTGTGAG 600

Qy    712 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCCTTTCTATTCGGAGCGGCGGCTGAAGGAG 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCCTTTCTATTCGGAGCGGCGGCTGAAGGAG 660

Qy    772 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACA 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACA 720

Qy    832 GTCACCGACAACATGCTGT-GTGCTGAGACACTCGG--AGCGGCGGCCCAGGCAAAC 888
              ||||||||||||||||||  ||||||||||||||||   |||||||||||||||
Db    721 GTCACCGACAACATGCTGTGGTGCTGGAGACACTCGGGAGGACGGCGGCCCCCAGGCAAAC 780

Qy    889 TTGCACGACGCCT-GCCAGGGGGGATTCGGGAGGCCCCCTGGT-GTGTCTGAAGGATGG-C 945
              |||||||||||   ||||||||||||||||||||||||||   |||||||||||||| 
Db    781 TTGCACGACGCCTGGCCGGGGGGGCGATTCGGGAGGCCCCCTGGTGGGTCTGAACGCATGGCC 840

Qy    946 CGGATGACTTT--GGTGGGCATCATCAGCTGGGGCCTCGGGCTGTGG 989
              |||||||||   |||||||||||||||||||||||||||||||||
Db    841 CGGATGACTTTTGGGGGGGCATCATCAACTGGGGGCCTCGGGCTGTGG 886
```

```
RESULT 10
BI765247
LOCUS       BI765247               916 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603050338F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190656 5',
            mRNA sequence.
ACCESSION   BI765247
VERSION     BI765247.1  GI:15755825
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 916)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11476 row: k  column: 09
            High quality sequence stop: 837.
FEATURES    Location/Qualifiers
source      1..916
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5190656"
            /lab_host="DH10B"
            /clone_lib="NIH MGC 116"
            /note="Organ: pooled colon, kidney, stomach; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning).  Average insert size 1.4 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
```

```
BASE COUNT      204 a    272 c    256 g    184 t
ORIGIN          Gruber (Invitrogen). Research Genetics tracking code
                023. Note: this is a NIH_MGC Library."

Query Match          72.1%;   Score 767.4;  DB 12;  Length 916;
Best Local Similarity 99.7%;   Pred. No. 1.3e-189;
Matches  779;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   1;

Qy   286 TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTTCCCACCCC 345
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTTCCCACCCC 60

Qy   346 TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCCGTTCCTGTGCGGG 405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCCGTTCCTGTGCGGG 120

Qy   406 GGCATACTCATCAGCTCCGTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTT 465
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGCATACTCATCAGCTCCGTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTT 180

Qy   466 CCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACGGGTGGTCCTGGCGGAGGAG 525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACGGGTGGTCCTGGCGGAGGAG 240

Qy   526 GAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTAC 585
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTAC 300

Qy   586 GACAATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGAGC 645
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GACAATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGAGC 360

Qy   646 AGCGTGGTCCGACACTGTGTGCCTTCCCCGGCGGACTGCAGCTGCCGGACTGGACGGAG 705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGCGTGGTCCGACACTGTGTGCCTTCCCCGGCGGACTGCAGCTGCCGGACTGGACGGAG 420

Qy   706 TGTGAGCCTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTG 765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TGTGAGCCTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTG 480

Qy   766 AAGGAGGCTCATGTGACGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAAC 825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGGAGGCTCATGTGACGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAAC 540

Qy   826 AGAACAGTCACCGACACATGCTGTGTCTGAGACACTCGGAGCGCGGCGCCCCAGGCA 885
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AGAACAGTCACCGACACATGCTGTGTCTGAGACACTCGGAGCGCGGCGCCCCAGGCA 600

Qy   886 AACTTGCACGACGCGCCTGCCAGGGGCGATTCGGAGGGCCCCTCGGTGTCTGAACGATGGC 945
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AACTTGCACGACGCGCCTGCCAGGGGCGATTCGGAGGGCCCCTCGGTGTCTGAACGATGGC 660

Qy   946 CGGATGACTTTGGTGGGCATCATCAGCTGGGGCTGTGTGACAGAAGGAGATGTCCCG 1005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGGATGACTTTGGTGGGCATCATCAGCTGGGGCTGTGTGACAGAAGGAGATGTCCCG 720

Qy  1006 GGTGTGTGACACAAAGGTTACCAACTACCTAGAC-TGGATTCGTGACAACATGCGACCGTG 1064
             ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   721 GGTGTGTGACACAAAGGTTACCAACTACCTAGACTTGGATTCGTGACAACATGCGACCGTG 780

Qy  1065 A 1065
           |
Db   781 A 781

RESULT 11
BQ927899
LOCUS       BQ927899     8775988 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6373489
                AGENCOURT_8775988 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6373489
                5', mRNA sequence.                            950 bp  mRNA  linear  EST 20-AUG-2002
DEFINITION
ACCESSION   BQ927899
VERSION     BQ927899.1  GI:22342930
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 950)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2550  row: p  column: 02
            High quality sequence start: 35
            High quality sequence stop: 684.
FEATURES             Location/Qualifiers
     source          1..950
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6373489"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dr priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT       207 a    283 c    279 g    181 t
ORIGIN

Query Match          71.8%;   Score 765.2;  DB 13;  Length 950;
Best Local Similarity 98.1%;   Pred. No. 4.9e-189;
Matches  806;  Conservative   0;  Mismatches  13;  Indels   3;  Gaps   3;

Qy   210 CCACGTGCTGAAGAACCGCAGCGTGACGTGGGAGTACGTGATGTGCCCTCCTGCTCCAC 269
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43 CCCCGTGCTGAAGAACCGCAGCGTGACGTGGGAGTACGTGATGTGCCCTCCTGCTCCCAC 102

Qy   270 CTGCGGGCCTGAGACAGTACAGCCAGCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGA 329
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   103 CTGCGGGCCTGAGACAGTACAGCCAGCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGA 162

Qy   330 CATCGCCTTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGA 389
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   163 CATCGCCTTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGA 222

Qy   390 GCCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCGTGGATTCTCTCTGCCGCCCACTG 449
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   223 GCCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCGTGGATTCTCTCTGCCGCCCACTG 282

Qy   450 CTTCCAGGAGAG-TTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGG 508
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   283 CTTCCAGGAGAGAGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGG 342

Qy   509 TGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAAT 568
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 TGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAAT 402

Qy   569 TCCATGATGACACTTACGACAATGACAATGACATTGCGCTGCTGCAGTCGAAATCGGATTCGTCCC 628
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   403 TCCATGATGACACTTACGACAATGACAATGACATTGCGCTGCTGCAGTCGAAATCGGATTCGTCCC 462
```

BASE COUNT   173 a   295 c   293 g   187 t
ORIGIN

Query Match   66.4%;   Score 706.8;   DB 10;   Length 948;
Best Local Similarity   93.3%;   Pred. No. 8.9e-174;
Matches 793;   Conservative 0;   Mismatches 52;   Indels 5;   Gaps 5;

```
Qy      6 GGGAAACAGTGACTGCTACT-TTGGGAATGGGTCAGCTCAGCTGGCACGCACAGCCTCA 64
            ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db      8 GAGGAACAGTGACTGCTATTATTGGGAATGGGTCAGCTCAGCTGGCACGCACAGCCTCA 67

Qy     65 CCGAGTCGGGTGCCTCCTGCCTCC-GTGGAATTCCATGATCCTGATAGGCAAGGTTTAC 123
            ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db     68 CCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTAC 127

Qy    124 ACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAAT 183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 ACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAAT 187

Qy    184 CCTGATGGGGATGCCAAGCCTGGTGCCACGTGCTGCTGAAGAACCGCAGGTGACGTGGAG 243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 CCTGATGGGGATGCCAAGCCTGGTGCCACGTGCTGCTGAAGAACCGCAGGTGACGTGGAG 247

Qy    244 TACTGTGATGTGCCTCCTGTCACCTGCGGGCTGAGACAGTACAGCCAGCCTCAGTT 303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TACTGTGATGTGCCTCCTGTCACCTGCGGGCTGAGACAGTACAGCCAGCCTCAGTT 307

Qy    304 CGCATCAAAGAGAGGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATCTT 363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 CGCATCAAAGAGAGGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATCTT 367

Qy    364 GCCAAGCACAGAGAGTCGCCCCGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCC 423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 GCCAAGCACAGAGAGTCGCCCCGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCC 427

Qy    424 TGCTGGAATTCTCTCGCCGCGCCCACTGCTTCGCCCACCACCTGACG 483
            |||||||||||||||||||||||||||||||||||||||||||||||
Db    428 TGCTGGAATTCTCTCGCCGCGCCCACTGCTTCGCCCACCACCTGACG 487

Qy    484 GTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTC 543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 GTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTC 547

Qy    544 GAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACAATGACATTGCCGCTG 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACAATGACATTGCCGCTG 607

Qy    604 CTGCAGTCGAAATCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGACTGTG 663
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 CTGCAGTCGAAATCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGACTGTG 667

Qy    664 TGCCTTCCCCGCGGGAACCTGCAGCTGCCGGACTGGACGGTGTGAGCTCTCCGGCTAC 723
            ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 TGCCTT-CCCCGCGGGAACCTGCAGCTGCCGGACTGGACGGTGTGAGCTCTCGGCTAC 726

Qy    724 GGCAAGCATGAGGAGGCCTTGTCCTTTCATTCGGAGGGGCTGAAGAGGCTCATGTCAGA 783
            ||||||||||||||||| ||| |||||||||||||||||||||||||| | |||||||
Db    727 GGCAAGCATGAGG-CTTGTCTCCTTTCATTCGGAGGGGCTGAAGAGGCTCATGTCGGA 785

Qy    784 CTGTACCCTCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACGACAAC 843
            |||     |||| || | ||||||||| | |||   | | ||||||||||||||||
Db    786 CTG-GTCCATCCAGCCCGCTGGGCTTCCCCACCTTTCTTCAGGCACAGCAGCCGCCTGC 844

Qy    844 ATGCTGTGTG 853
            |        |
Db    845 TGGGGCTGGG 854
```

RESULT 12
BE616613
LOCUS   BE616613   948 bp   mRNA   linear   EST 24-AUG-2000
DEFINITION 601278864F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3611007 5',
            mRNA sequence.
ACCESSION   BE616613
VERSION   BE616613.1 GI:9899212
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 948)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM267 row: p column: 16
            High quality sequence start: 2
            High quality sequence stop: 759.
FEATURES   Location/Qualifiers
     source   1..948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3611007"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_39"
                     /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
                     Site 2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCCACGAG(G). Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California, Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life

RESULT 13
BG763582
LOCUS   BG763582   958 bp   mRNA   linear   EST 15-MAY-2001
DEFINITION 602735987F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4861268 5',

mRNA sequence.

ACCESSION   BG763582
VERSION   BG763582.1   GI:14074235
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 958)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1719 row: n column: 21
            High quality sequence stop: 810.
FEATURES   Location/Qualifiers
     source   1..958
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4861268"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_49"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."

BASE COUNT   198 a   285 c   282 g   193 t
ORIGIN

Query Match   65.5%;   Score 697.4; DB 12; Length 958;
Best Local Similarity 94.2%; Pred. No. 2.6e-171;
Matches   834; Conservative   0; Mismatches 36; Indels 15; Gaps 10;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||| |||||||||||||||||||||||||||||||||||||||| ||||||||
Db   72 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 131

Qy   61 CTCACCGAGTCGGGTGGCCTCCTGCCTCCCGTGCCTCCGTGGAATTCCATGATCCTGAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132 CTCACCGAGTCGGGTGGCCTCCTGCCTCCCGTGCCTCCGTGGAATTCCATGATCCTGAAG 191

Qy   121 TACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGGG 251

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCTGAGAACCGAGGCTGACGTGG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 AATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCTGAGAACCGAGGCTGACGTGG 311

Qy   241 GAGTACTGTGATGTGCCCTCCTTCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   312 GAGTACTGTGATGTGCCCTCCTTCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 371

Qy   301 TTTCGCATCAAAGGAGAGGGCTCTTCGCCGACATCGCCCCCACCCCTGGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 TTTCGCATCAAAGGAGAGGGCTCTTCGCCGACATCGCCCCCACCCCTGGCAGGCTGCCATC 431

Qy   361 TTTGCCAAGCACCAGCAGGAGTCGCCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420

Db   432 TTTGCCAAGCACCAGCAGGAGTCGCCGCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 491
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 TCCTGCTGGATTCTTCTGCGCCCCACTGCTTCCAGGAGGAGGTTTCGCCCCCACCACCTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   492 TCCTGCTGGATTCTTCTGCGCCCCACTGCTTCCAGGAGGAGGTTTCCTGCCCACCACCTG 551

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTGAA 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTGAA 611

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACATTGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   612 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGACATTGCG 671

Qy   601 CTGCTGCAGCTGAAATGGATTCGT--CCCGCTGTGCCCAGGAGCAGC-GTGGTCCGCA 658
              |||||||||||||||||||||||| |||||||||||||||||||||| |||||||||
Db   672 CTGCTGCAGCTGAAATCGGATTCGTCCCCGCTGTGCCCAGGAGCAGCGGGTGTCCGCA 731

Qy   659 CGTGTGTGCCTT--CCCCGGCGACCTGCAGCTGCCGGACTGGAC-GGAGTGTGAGCTCT 715
              |||||||||||| |||||||||||||||||||||||||||| |||||||||||||||
Db   732 CGTGTGTGCCTTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGGAGTGTGAGCTCT 791

Qy   716 CCG-GCTACGGCAAGCAT--GAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGG 772
              ||| |||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db   792 CCGAGCTACGGCAAGCATTTGAGGGCCTGGTCTCCTTTCTATTCGGAGGCGG-TGAAGGAG 850

Qy   773 CTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTACAGAACAG 832
              |||||| |||||| || ||| || |||||| ||| |||||||||||||||||||
Db   851 CTCATG--CGACTGGACCCCATTCCACCCGGTGGCCATTCCACAACTTTACTT-ACAGAACAG 907

Qy   833 TCACCGACAACATGCTGTGTGTGGAGACACTCGGAGCGCGCGGC 877
              |||||||||||||||||||||||||||||||||||||||||||
Db   908 TACC--GAAAATGGTGTGTGCGGGACCCCGACGGGGGCCCAGGC 949

RESULT 14
BQ960216
LOCUS   BQ960216   974 bp   mRNA   linear   EST 21-AUG-2002
DEFINITION AGENCOURT_8945604 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6459629
            5', mRNA sequence.
ACCESSION   BQ960216
VERSION   BQ960216.1   GI:22375694
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 974)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2645 row: a   column: 06
            High quality sequence stop: 603.
FEATURES   Location/Qualifiers
     source   1..974
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6459629"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_101"
                     /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:

```
VERSION     BE732704.1  GI:10146696
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 709)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM752  row: d  column: 17
            High quality sequence start: 3
            High quality sequence stop: 701.
FEATURES             Location/Qualifiers
     source          1..709
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3925600"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_21"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT    147 a   219 c   203 g   140 t
ORIGIN
```

Query Match   65.1%;   Score 693.4; DB 10;   Length 709;
Best Local Similarity 99.7%;   Pred. No. 2.5e-170;
Matches 705;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 1;

```
Qy   197 CCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTGC 256
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTGC 60

Qy   257 CCTCCTGCTCCACCTGCGGCCTGAGACACTGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAG 316
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CCTCCTGCTCCACCTGCGGCCTGAGACACTGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAG 120

Qy   317 GGCTCTTCGCGGACATCGCCTCCACCCCTGGCGGGGTGCCATCTTTGCCAAGCACAGGA 376
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGCTCTTCGCGGACATCGCCTCCACCCCTGGCGGGGTGCCATCTTTGCCAAGCACAGGA 180

Qy   377 GGTGCGCCCGGAGAGCGGTTCCTGTGCGGGGGACATACTCATCAGCTCCTGCTTCTCT 436
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGTGCGCCCGGAGAGCGGTTCCTGTGCGGGGGACATACTCATCAGCTCCTGCTTCTCT 240

Qy   437 CTGCCGCGCCACTGCTTCCAGGAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCA 496
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGCCGCGCCACTGCTTCCAGGAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCA 300

Qy   497 GAACATACCGGGTGCCCGGAGGAGGAGAATTTGAAGTCAAAAATACATTG 556
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAACATACCGGGTGCCCGGAGGAGGAGAATTTGAAGTCAAAAATACATTG 360

Qy   557 TCCATAAGGAATTCGATGATGACACTTACGACAATGACAGTGCGCTGCTGCAGCTGAAT 616
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCATAAGGAATTCGATGATGACACTTACGACAATGACAGTGCGCTGCTGCAGCTGAAT 420
```

Wed Aug 20 09:29:44 2003

```
Qy   617 CGGATTGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCCGG  676
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGGATTCGTCCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCCGG  480

Qy   677 CGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTTCCGGCTACGGCAAGCATGAGG  736
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTTCCGGCTACGGCAAGCATGAGG  540

Qy   737 CCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGAGGGCTCATGTCAGACTGTACCATCCA  796
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTTGTCCCCTTTCTATTCGGAGCGGCTGAAGAGGGCTCATGTCAGACTGTACCCATCCA  600

Qy   797 GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG  856
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG  660

Qy   857 GAGACACTCGGAGCGGGGCGCGGGCCCCAGGCAAACTTGCACGACGCCTGC  903
         |||||||||||||||||||||||| |||||||||||||||||||||||||
Db   661 GAGACACTCGGAGCGCGGCGGGG-CCCAGGCAAACTTGCACGACGCCTGC  706
```

Search completed: August 19, 2003, 08:29:20
Job time : 3116.64 secs

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 19:53:48 ; Search time 4471.78 Seconds
                                              (without alignments)
                                              9743.047 Million cell updates/sec

Title: US-09-987-455-4
Perfect score: 1065
Sequence: 1 tctgagggaaacagtgactg........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :   GenEmbl:*
                     1: gb_ba:*
                     2: gb_htg:*
                     3: gb_in:*
                     4: gb_om:*
                     5: gb_ov:*
                     6: gb_pat:*
                     7: gb_ph:*
                     8: gb_pl:*
                     9: gb_pr:*
                    10: gb_ro:*
                    11: gb_sts:*
                    12: gb_sy:*
                    13: gb_un:*
                    14: gb_vi:*
                    15: em_ba:*
                    16: em_fun:*
                    17: em_hum:*
                    18: em_in:*
                    19: em_mu:*
                    20: em_om:*
                    21: em_or:*
                    22: em_ov:*
                    23: em_pat:*
                    24: em_ph:*
                    25: em_pl:*
                    26: em_ro:*
                    27: em_sts:*
                    28: em_un:*
                    29: em_vi:*
                    30: em_htg_hum:*
                    31: em_htg_inv:*
                    32: em_htg_other:*
                    33: em_htg_mus:*
                    34: em_htg_pln:*
                    35: em_htg_rod:*
                    36: em_htg_mam:*
                    37: em_htg_vrt:*
                    38: em_sy:*
                    39: em_htgo_hum:*
                    40: em_htgo_mus:*
                    41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1065	100.0	1065	6	AX431270	AX431270 Sequence
2	1065	100.0	1065	6	AX431273	AX431273 Sequence
3	1065	100.0	1068	6	A27431	A27431 DNA sequenc
4	1065	100.0	1068	6	AR059987	AR059987 Sequence
5	1065	100.0	1068	6	E01935	E01935 Synthetic D
6	1065	100.0	1170	6	A27727	A27727 DNA sequenc
7	1065	100.0	1170	6	E01945	E01945 Synthetic D
8	1065	100.0	1314	6	A27435	A27435 DNA sequenc
9	1065	100.0	1314	6	E01937	E01937 Synthetic D
10	1065	100.0	1419	6	A27725	A27725 DNA sequenc
11	1065	100.0	1419	6	AR059996	AR059996 Sequence
12	1065	100.0	1419	6	E01944	E01944 Synthetic D
13	1065	100.0	1689	6	I06609	I06609 Sequence 44
14	1065	100.0	1689	6	I06614	I06614 Sequence 46
15	1065	100.0	1689	6	I08789	I08789 Sequence 3
16	1065	100.0	1780	6	A06611	A06611 Synthetic n
17	1065	100.0	1780	6	A10226	A10226 Synthetic D
18	1065	100.0	1836	6	E01176	E01176 DNA encodin
19	1065	100.0	1836	6	I01581	I01581 Sequence 1
20	1065	100.0	1981	6	I01583	I01583 Sequence 3
21	1065	100.0	1989	6	I01256	I01256 Sequence 2
22	1065	100.0	2099	6	A07207	A07207 Artificial
23	1065	100.0	2100	6	A07197	A07197 H.sapiens m
24	1065	100.0	2100	6	A27429	A27429 cDNA sequen
25	1065	100.0	2100	6	E01934	E01934 DNA encodin
26	1065	100.0	2101	6	AR059986	AR059986 Sequence
27	1065	100.0	2151	6	E00654	E00654 cDNA encodi
28	1065	100.0	2162	6	AR017909	AR017909 Sequence
29	1065	100.0	2162	6	AR017931	AR017931 Sequence
30	1065	100.0	2162	6	I18796	I18796 Sequence 3
31	1065	100.0	2249	6	A07209	A07209 Artificial
32	1065	100.0	2544	6	I08530	I08530 Sequence 1
33	1065	100.0	2547	6	E01163	E01163 cDNA encodi
34	1065	100.0	2547	6	I07841	I07841 Sequence 1
35	1065	100.0	2547	6	I07989	I07989 Sequence 1
36	1065	100.0	2547	6	I09622	I09622 Sequence 1
37	1065	100.0	2549	6	E01055	E01055 DNA encodin
38	1065	100.0	2560	6	A01465	A01465 H.sapiens t
39	1065	100.0	7533	6	I08638	I08638 Sequence 2
40	1064	99.9	1689	6	I08637	I08637 Sequence 1
41	1063.4	99.8	1170	6	AR059997	AR059997 Sequence
42	1063.4	99.8	1314	6	AR059989	AR059989 Sequence
43	1063.4	99.8	1506	6	E02247	E02247 cDNA sequen
44	1063.4	99.8	1680	6	I06422	I06422 Sequence 40
45	1063.4	99.8	1689	6	A14927	A14927 tPA-DNA fro

ALIGNMENTS

RESULT 1
AX431270
LOCUS   AX431270   1065 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 4 from Patent WO0240650.
ACCESSION   AX431270
VERSION   AX431270.1 GI:21656152
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 4 23-MAY-2002;

Pred. No. is the number of results predicted by chance to have a

BOEHRINGER INGELHEIM INT (DE)

```
FEATURES      Location/Qualifiers
  source      1..1065
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
              /note="coding sequence for K2S protein"
BASE COUNT    225 a    314 c    312 g    214 t
ORIGIN

Query Match          100.0%;  Score 1065;  DB 6;  Length 1065;
Best Local Similarity 100.0%;  Pred. No. 4.7e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60

Qy     61 CTCACCCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCCCGTGCTGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCACCCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCCCGTGCTGATCCTGATAGGCAAGGTT 120

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGCCACCCCAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACAGCACAGAACCCCAGTGCCCAGCCACCCCAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGG 180

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGACCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGACCAGGCTGACGTGG 240

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACTCGCGGCCTGAGACAGCAGTACAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGTACTGTGATGTGCCCTCCTGCTCCACTCGCGGCCTGAGACAGCAGTACAGCCAGCCTCAG 300

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC 360

Qy    361 TTTGCCAAGCACAGGAGAGTCGCCCGAGAGCCGGTTCGCCCCGGAGAGGAGTTTCCGCCGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTGCCAAGCACAGGAGAGTCGCCCGAGAGCCGGTTCGCCCCGGAGAGGAGTTTCCGCCGG 420

Qy    421 TCCTGCTGGATTCTCTCGCCGCCGGACTGCCAGCTGCCAGACTGGACCGGAGTGGACCCCGGACTCTCCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCTGCTGGATTCTCTCGCCGCCGGACTGCCAGCTGCCAGACTGGACCGGAGTGGACCCCGGACTCTCCGGC 480

Qy    481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGACAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGACAGAAATTTGAA 540

Qy    541 GTCGAAAAATACATTGTCTCATAAGGAATTCGATGAGATGACACTTACGACACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCGAAAAATACATTGTCTCATAAGGAATTCGATGAGATGACACTTACGACACAATGACATTGCG 600

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGGGTGTCCGCTGTGCCAGGAGAGCAGCGCGGCGACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGGGTGTCCGCTGTGCCAGGAGAGCAGCGCGGCGACT 660

Qy    661 GTGTGCCAAGCAAGCCTTGACCAGCTGCCGGACTGCCGAGTGTGACGGCGCTGAGCTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTGTGCCAAGCAAGCCTTGACCAGCTGCCGGACTGCCGAGTGTGACGGCGCTGAGCTCTCCGGC 720

Qy    721 TACGGCAAGCATGAGGCCTTGTCTTCATTCGCACATCACAACATTACTTAACAGAACAGTCACGAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TACGGCAAGCATGAGGCCTTGTCTTCATTCGCACATCACAACATTACTTAACAGAACAGTCACGAC 780

Qy    781 AGACTGTACCCATCCAGCCGCGTCGACATCCAGCCGCGTCGACATTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGACTGTACCCATCCAGCCGCGTCGACATCCAGCCGCGTCGACATTACTTAACAGAACAGTCACCGAC 840

Qy    841 AACATGCTGTGTGTGCTGGAGACACTCGGAGACTGTGTGCTGGAGACACTCGGAGACACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACATGCTGTGTGTGCTGGAGACACTCGGAGACTGTGTGCTGGAGACACTCGGAGACACTTGCACGACGCC 900
```

```
Qy    901 TGCCAGGGGCGATTCGGAGGGCCCCCTGGTGTCGTCTGAACGATGGCCGATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TGCCAGGGGCGATTCGGAGGGCCCCCTGGTGTCGTCTGAACGATGGCCGATGACTTTGGTG 960

Qy    961 GGCATCATCAGCTGGGGGCCTGGGGCTGGGGCCTGGGGCGTGTGGACAGAAGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCATCATCAGCTGGGGGCCTGGGGCTGGGGCCTGGGGCGTGTGGACAGAAGATGTCCCGGGTGTGTACACAAAG 1020

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
```

```
RESULT 2
AX431273
LOCUS       AX431273              1065 bp   DNA       linear   PAT 28-JUN-2002
DEFINITION  Sequence 7 from Patent WO0240650.
ACCESSION   AX431273
VERSION     AX431273.1  GI:21656155
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE     Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 7 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES      Location/Qualifiers
  source      1..1065
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
              /note="coding sequence for K2S protein"
BASE COUNT    225 a    314 c    312 g    214 t
ORIGIN

Query Match          100.0%;  Score 1065;  DB 6;  Length 1065;
Best Local Similarity 100.0%;  Pred. No. 4.7e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60

Qy     61 CTCACCCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCCCGTGCTGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCACCCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCCCGTGCTGATCCTGATAGGCAAGGTT 120

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGCCACCCCAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACAGCACAGAACCCCAGTGCCCAGCCACCCCAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGG 180

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGACCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGACCAGGCTGACGTGG 240

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACTCGCGGCCTGAGACAGCAGTACAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGTACTGTGATGTGCCCTCCTGCTCCACTCGCGGCCTGAGACAGCAGTACAGCCAGCCTCAG 300

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC 360

Qy    361 TTTGCCAAGCACAGGAGAGTCGCCCGAGAGCCGGTTCGCCCCGGAGAGGAGTTTCCGCCGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTGCCAAGCACAGGAGAGTCGCCCGAGAGCCGGTTCGCCCCGGAGAGGAGTTTCCGCCGG 420

Qy    421 TCCTGCTGGATTCTCTCGCCGCCGGACTGCCAGCTGCCAGACTGGACCGGAGTGGACCCCGGACTCTCCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCTGCTGGATTCTCTCGCCGCCGGACTGCCAGCTGCCAGACTGGACCGGAGTGGACCCCGGACTCTCCGGC 480
```

```
Db    4  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGCACGCACCGCACCGCAGC   63

Qy   61  CTCACCGAGTCGGGTGCGTGCCTCCTGCTCCCGCTCCCGCCCGTGGGCGGAATTCCATGATCCGTGAATCCTGATAGGCAAGGTT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64  CTCACCGAGTCGGGTGCGTGCCTCCTGCTCCCGCTCCCGCCCGTGGGCGGAATTCCATGATCCGTGAATCCTGATAGGCAAGGTT  123

Qy  121  TACACAGCAGAACCCAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCGG  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124  TACACAGCAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGG  183

Qy  181  AATCCTGATGGGGATGCCAAGCCTGCTGCACGTGCTGAAGAACCGAGGCTGACGTGG  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184  AATCCTGATGGGGATGCCAAGCCTGCTGCACGTGCTGAAGAACCGAGGCTGACGTGG  243

Qy  241  GAGTACTGTGATGTGCCCCTCGCCTGCAGCTGCCCTGAGACTACGCCAGCCTCAG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244  GAGTACTGTGATGTGCCCCTCGCCTGCAGCTGCCCTGAGACTACGCCAGCCTCAG  303

Qy  301  TTTCGCATCAAAGGAGGGGCTCTTCGCCACGACGCCCCTGGCAGGCTGCCATC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  304  TTTCGCATCAAAGGAGGGGCTCTTCGCCACGACGCCCCTGGCAGGCTGCCATC  363

Qy  361  TTTGCCAAGCACAGGAGTCGCCCGAGAGCGGTTCCTGCGTGCGGGGGCATACTCATCAGC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364  TTTGCCAAGCACAGGAGTCGCCCGAGAGCGGTTCCTGCGTGCGGGGGCATACTCATCAGC  423

Qy  421  TCCTGCTGATTCTCTGCGGGACACTCGGAGGTTTCCGCCCCACCACCTG  480
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  424  TCCTGCTGATTCTCTGCGGGACACTCGGAGGTTTCCGCCCCACCACCTG  483

Qy  481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA  543

Qy  541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  603

Qy  601  CTGCTCGAGCTGAAATCGATTCGTCCCCGTGTCCCCGACTCGACGAGCAGCGTGGTCCGCACT  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604  CTGCTCGAGCTGAAATCGATTCGTCCCCGTGTCCCCGACTCGACGAGCAGCGTGGTCCGCACT  663

Qy  661  GTGTGCCTTCCCCGGCCGGACCTGCAGCTGACGACTGAGCGTGAGCGTCTCCGGC  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664  GTGTGCCTTCCCCGGCCGGACCTGCAGCTGACGACTGAGCGTGAGCGTCTCCGGC  723

Qy  721  TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  724  TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  783

Qy  781  AGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784  AGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  843

Qy  841  AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGCCCCAGGCCCAAACTTGCACGACGCC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  844  AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGCCCCAGGCCCAAACTTGCACGACGCC  903

Qy  901  TGCCAGGGCGATTCGGGAGGCCCCCGTGGTGTCTGAACGATGCCGCATGACTTTGGTG  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  904  TGCCAGGGCGATTCGGGAGGCCCCCGTGGTGTCTGAACGATGCCGCATGACTTTGGTG  963

Qy  961  GGCATCATCAGCTGGGGCTGGGCGTGTGTGACAGAGATGTCCCGGGTGTGTACACAAAG  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  964  GGCATCATCAGCTGGGGCTGGGCGTGTGTGACAGAGATGTCCCGGGTGTGTACACAAAG  1023

Qy 1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA  1065
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 1024  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA  1068
```

Query Match   100.0%; Score 1065; DB 6; Length 1068;
Best Local Similarity   100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGCACGCACCGCACCGCAGC   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    724 TACGGCAAGCATGAGGCCTTGCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 783
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          |||||||||  |||||||||||||||||||||  |||||  |||||||||||||||||||
Db    784 AGACTGTACCCATCCAGCCGCTGCACATCACACAACATTTAACAGAACAGTCACCGAC 843
Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGCCGGCCCCAGGCAAACTTGCACGACGCC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGGCCCCAGGCAAACTTGCACGACGCC 903
Qy    901 TGCCAGGCGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGCCGATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 TGCCAGGGCGGATTCGGGAGGCCCCCCTGGTGTGTCTGAACGATGCCGATGACTTTGGTG 963
Qy    961 GGCATCATCAGCTGGGGCCTGGGGCCTGGCGTGCTGGACAAGAAGGATGTCCCGGGTGTGTACCACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 GGCATCATCAGCTGGGGCCTGGGGCCTGGCGTGCTGGACAAGAAGGATGTCCCGGGTGTGTACCACAAAG 1023
Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1068
```

RESULT 5
E01935

```
LOCUS       E01935                  1068 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Synthetic DNA encoding new tissue plasminogen activator (t-PA).
ACCESSION   E01935
VERSION     E01935.1  GI:2170183
KEYWORDS    JP 1989104167-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 2 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/2
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87  8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19)
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   *source: clone=pTTkPAdeltatrp;
            FH   Key             Location/Qualifiers
            FH
            FT   CDS             1..1068
            FT                   /product='t-PA'.
FEATURES             Location/Qualifiers
     source          1..1068
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      226 a    314 c    313 g    215 t
ORIGIN

Query Match           100.0%;  Score 1065;  DB 6;  Length 1068;
Best Local Similarity 100.0%;  Pred. No. 4.7e-248;
Matches 1065;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCCACAGC 60
```

```
LOCUS       AR059987              1068 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 44 from patent US 5840533.
ACCESSION   AR059987
VERSION     AR059987.1  GI:5986437
KEYWORDS    Unknown.
SOURCE      Unknown.
  ORGANISM  Unclassified.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 44 24-NOV-1998;.
FEATURES             Location/Qualifiers
     source          1..1068
                     /organism="unknown"
BASE COUNT      226 a    314 c    313 g    215 t
ORIGIN

Query Match           100.0%;  Score 1065;  DB 6;  Length 1068;
Best Local Similarity 100.0%;  Pred. No. 4.7e-248;
Matches 1065;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCCACAGC 63
Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTCCCGTGGAATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTCCCGTGGAATCCTGATAGGCAAGGTT 123
Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCTGGCCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCTGGCCAAACATAATTACTGCCGG 183
Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 243
Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCCTGAGACAGTAGCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCCTGAGACAGTAGCAGCCTCAG 303
Qy    301 TTTCGCATCAAACGAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 TTTCGCATCAAACGAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 363
Qy    361 TTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 423
Qy    421 TCCTGCTGGATTCTCTGCGCCACTGGGTTCCCGCCCCACCACTG 480
          |||||||||||||||||||||||||||||||||||||||||||
Db    424 TCCTGCTGGATTCTCTGCGCCACTGGGTTCCCGCCCCACCACTG 483
Qy    481 ACGGTGATCTTGGGCAGAACATACGGGGTGTCCTGGCGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 ACGGTGATCTTGGGCAGAACATACGGGGTGTCCTGGCGAGGAGCAGAAATTTGAA 543
Qy    541 GTCGAAAAATACATTGTGTCCATAAGGAATTCGATGATGATGACACTTACGACGACAATGGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 GTCGAAAAATACATTGTGTCCATAAGGAATTCGATGATGATGACACTTACGACGACAATGGCG 603
Qy    601 CTGCTGCGAGCTGAAATCGGATTCGTCCCGTGTCCCGCTGCGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 CTGCTGCGAGCTGAAATCGGATTCGTCCCGTGTCCCGCTGCGTGGTCCGCACT 663
Qy    661 GTGTGCCTTCCCCGGCGGACTGCAGCTGCCGGACTGCCGAGTGTGAGCGTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 GTGTGCCTTCCCCGGCGGACTGCAGCTGCCGGACTGCCGAGTGTGAGCGTCTCCGGC 723
Qy    721 TACGGCAAGCATGAGGCCTTGCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
A27727
LOCUS         A27727                          1170 bp    DNA              linear   PAT 09-JUL-2002
DEFINITION    DNA sequence of coding region in pmTTk.
ACCESSION     A27727
VERSION       A27727.1  GI:2172727237
KEYWORDS      .
SOURCE        synthetic construct
  ORGANISM    synthetic construct
              artificial sequences.
REFERENCE     1 (bases 1 to 1170)
  AUTHORS     Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
              Kobayashi.M.
  TITLE       New tissue plasminogen activator
  JOURNAL     Patent: EP 0302456-A 51 08-FEB-1989;
              FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1170
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        249 a      338 c      346 g      237 t
ORIGIN
```

Query Match   100.0%; Score 1065; DB 6; Length 1170;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy       1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC  165

Qy      61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  225

Qy     121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGAAGAACATAATTACTGCGCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     226 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGAAGAACATAATTACTGCGCGG  285

Qy     181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  345

Qy     241 GAGTACTGTGATGTGCCCTGTGCCCTCGCGACTGTGAGCTCTCCGGC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     346 GAGTACTGTGATGTGCCCTGTGCCCTCGCGACTGTGAGCTCTCCGGC  405

Qy     301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC  465

Qy     361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGGCATACTACTCAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGGCATACTACTCAGC  525

Qy     421 TCCTGCTGGATTCTCTGCTCTTCCAGGAGGAGTTCCGCCCACCACCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     526 TCCTGCTGGATTCTCTGCTCTTCCAGGAGGAGTTCCGCCCACCACCTG  585

Qy     481 ACGGTGATCTTGGGCAGACATACCGGGTGGTCTGCGGATGATGCAATGACATTGGTG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     586 ACGGTGATCTTGGGCAGACATACCGGGTGGTCTGCGGATGATGCAATGACATTGGTG  645

Qy     541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAGTCACAAAG  600
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     646 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAGTCACAAAG  705

Qy     601 CTGCTGCAGCTGAAATCGGATTCGTGACAACATGCAGCCGTGA  660
           |||||||||||||||||||||||||||||||||||||||||||
Db     706 CTGCTGCAGCTGAAATCGGATTCGTGACAACATGCAGCCGTGA  765

Qy     661 GTGTGCCTTCCCCGGCCAGCTGCAGCTGGACGGAGTGGACCTCTCCGGC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       4 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC   63
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  123

Qy     121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGAAGAACATAATTACTGCGCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGAAGAACATAATTACTGCGCGG  183

Qy     181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  243

Qy     241 GAGTACTGTGATGTGCCCTGTGCCCTCGCGACTGTGAGCTCTCCGGC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 GAGTACTGTGATGTGCCCTGTGCCCTCGCGACTGTGAGCTCTCCGGC  303

Qy     301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC  363

Qy     361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGGCATACTACTCAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGGCATACTACTCAGC  423

Qy     421 TCCTGCTGGATTCTCTGCTCTTCCAGGAGGAGTTTCCGCCCACCACCTG  480
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 TCCTGCTGGATTCTCTGCTCTTCCAGGAGGAGTTTCCGCCCACCACCTG  483

Qy     481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCTGCGGAGGAGCAGAAATTTGAA  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCTGCGGAGGAGCAGAAATTTGAA  543

Qy     541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAGTCACAAAG  600
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAGTCACAAAG  603

Qy     601 CTGCTGCAGCTGAAATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCGACT  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 CTGCTGCAGCTGAAATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCGACT  663

Qy     661 GTGTGCCTTCCCCCGGCCGCAGCTGCCGCAGCTGCCGGCGTGAGCTTCTCCGGC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 GTGTGCCTTCCCCCGGCCGCAGCTGCCGCAGCTGCCGGCGTGAGCTTCTCCGGC  723

Qy     721 TACGGCAAGCATGAGGGCCTTGTCCTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     724 TACGGCAAGCATGAGGGCCTTGTCCTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC  783

Qy     781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     784 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC  843

Qy     841 AACATGCTGTGTGGAGGGGCGGGGCCCCCCTGTGTGCTGAACGAGCTTTTGGTG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     844 AACATGCTGTGTGGAGGGGCGGGGCCCCCCTGTGTGCTGAACGAGCTTTTGGTG  903

Qy     901 TGCCAGGCGATTCGGGAGGGCCCCCCTGTGTGCTGAACGATGGCCGCATGGACTTTTGGTG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     904 TGCCAGGCGATTCGGGAGGGCCCCCCTGTGTGCTGAACGATGGCCGCATGGACTTTTGGTG  963

Qy     961 GGCATCATCAGCTGGGCGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTGTACACAAAG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     964 GGCATCATCAGCTGGGCGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTGTACACAAAG 1023

Qy    1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCAGCCGTGA 1065
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCAGCCGTGA 1068
```

RESULT 6

```
Query Match          100.0%;  Score 1065;  DB 6;  Length 1170;
Best Local Similarity 100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1  TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGCACGCACAGC     60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       106  TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGCACGCACAGC    165

Qy        61  CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT    120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       166  CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT    225

Qy       121  TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG    180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       226  TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG    285

Qy       181  AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG    240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       286  AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG    345

Qy       241  GAGTACTGTGATGTGCCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG    300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       346  GAGTACTGTGATGTGCCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG    405

Qy       301  TTTCGCATCAAAGAGGAGGGCTCTTGCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC    360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       406  TTTCGCATCAAAGAGGAGGGCTCTTGCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC    465

Qy       361  TTTGCCAAGCACAGGAGTCGCCGCGGAGAGGTCGCCCGGAGAGGTTTCGCCCCACCACTG    420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       466  TTTGCCAAGCACAGGAGTCGCCGCGGAGAGGTCGCCCGGAGAGGTTTCGCCCCACCACTG    525

Qy       421  TCCTGCTGGATTCTCTGCGCCACTGCTTCCAGGAGACTGGACTGAGCTGTGAGCTCTCCGGC    480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       526  TCCTGCTGGATTCTCTGCGCCACTGCTTCCAGGAGACTGGACTGAGCTGTGAGCTCTCCGGC    585

Qy       481  ACGGTGATCTTGGGCAGAACATACCGGGTGTCCTGGCAGGAGCAGAAATTTGAA    540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       586  ACGGTGATCTTGGGCAGAACATACCGGGTGTCCTGGCAGGAGCAGAAATTTGAA    645

Qy       541  GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG    600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       646  GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG    705

Qy       601  CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGCAGCAGTGGTCCGCACT    660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       706  CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGCAGCAGTGGTCCGCACT    765

Qy       661  GTGTGCCTTCCCCGCGGGACTGGACGTGACTGAGCTCTCCGGC    720
              ||||||||||||||||||||||||||||||||||||||||||
Db       766  GTGTGCCTTCCCCGCGGGACTGGACGTGACTGAGCTCTCCGGC    825

Qy       721  TACGGCAAGCATGAGGCCTTGCTCCTTTCTATTCGGAAGGAGGCTCATGTC    780
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       826  TACGGCAAGCATGAGGCCTTGCTCCTTTCTATTCGGAAGGAGGCTCATGTC    885

Qy       781  AGAACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC    840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       886  AGAACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC    945

Qy       841  AACATGCTGTGTCTGGAGACACTCGGAGCGGCGCCCCAGGCAAACTTGCACGAGCGCC    900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       946  AACATGCTGTGTCTGGAGACACTCGGAGCGGCGCCCCAGGCAAACTTGCACGAGCGCC    1005

Qy       901  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGACGATGGCCGATGACTTTGGTG    960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1006  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGACGATGGCCGATGACTTTGGTG    1065

Qy       961  GGCATCATCAGCTCGGGCCTGTGGCTGTGACAGGAAGGATGTCCCGGCGTGTACACAAAG    1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1066  GGCATCATCAGCTCGGGCCTGTGGCTGTGACAGGAAGGATGTCCCGGCGTGTACACAAAG    1125

Qy      1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGGCGACCGTGA    1065
              ||||||||||||||||||||||||||||||||||||||||||||||
Db      1126  GTTACCAACTACCTAGACTGGATTCGTGACAACATGGCGACCGTGA    1170
```

RESULT 7
E01945
LOCUS   E01945   1170 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION   Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01945
VERSION   E01945.1 GI:2170193
KEYWORDS   JP 198910 4167-A/12.
SOURCE   JP 198910 4167-A/12
             21-APR-1989
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1 (bases 1 to 1170)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
             Kobayashi,M.
  TITLE   NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 198910 4167-A 12 21-APR-1989;
             FUJISAWA PHARMACEUT CO LTD
COMMENT   OS   Artificial gene
             OC   Artificial sequence; Genes.
             PN   JP 198910 4167-A/12
             PD   21-APR-1989
             PF   01-AUG-1988 JP 1988192320
             PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87   8725052, PR
             13-NOV-1987 GB 87 872 6683
             PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
             PI   NOTANI JOJI,
             PI   KOBAYASHI MASAKAZU
             PC   C12N9/64,C12N1/20,C12N15/00,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
             C12R1:19),
             PC   (C12N9/64,C12R1:91);
             CC   strandedness: Double;
             CC   topology: Linear;
             CC   hypothetical: No;
             CC   anti-sense: No;
             CC   *source: clone=pmTTk;
             FH   Key   Location/Qualifiers
             FH
             FT   CDS   1..1170
             FT   /product='new t-PA'.
FEATURES   Location/Qualifiers
     source   1..1170
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   249 a   338 c   346 g   237 t
ORIGIN

```
Db   1126 GTTACCAACTACCTAGACTGGATTCGTGTGACAACATGCGACCGTGA 1170
          |||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 8
A27435

LOCUS   A27435   1314 bp   DNA   linear   PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pTQkPAdeltatrp.
ACCESSION   A27435
VERSION   A27435.1 GI:2172723 0
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 35 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES   Location/Qualifiers
     source   1..1314
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   286 a   393 c   386 c   249 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1314;
Best Local Similarity 100.0%;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC    60
         ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  250 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC  309

Qy   61 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  310 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  369

Qy  121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  370 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  429

Qy  181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGAGCGTGACGTGG  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  430 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGAGCGTGACGTGG  489

Qy  241 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGCGTCACGTACAGCAGCCAGCCTCAG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  490 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGCGTCACGTACAGCAGCCAGCCTCAG  549

Qy  301 TTTCGGCATCAAAGGAGGAGGGCTCTTCGCCGACATGCCCACCCCTGCCAGGCTGCCATC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  550 TTTCGGCATCAAAGGAGGAGGGCTCTTCGCCGACATGCCCACCCCTGCCAGGCTGCCATC  609

Qy  361 TTTGCCAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTCGTGCGGGGGCATACTCATCAGC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  610 TTTGCCAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTCGTGCGGGGGCATACTCATCAGC  669

Qy  421 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  670 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG  729

Qy  481 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  730 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA  789

Qy  541 GTGAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACACATTGCG  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  790 GTGAAAATACATTGTCCATAAGGAATTCGATGATGACAATGACACATTGCG  849

Qy  601 CTGCTGCAGCTGAAATCGGATTCCCCGTCCCCGTGCCCGTGCCCTGTGGTTCCGCCACT  660
```

```
Db  850 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCGACTGACGAGCAGCGAGAGCGAGCTCTCCGGC  909
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 GTGTGCCTTCCCCGGCCGACCTGCCAGCTGCCAGCTGACCTGCGACTGCGAGTGTGAGCTCTCCGGC  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  910 GTGTGCCTTCCCCGGCCGACCTGCCAGCTGCCAGCTGACCTGCGACTGCGAGTGTGAGCTCTCCGGC  969

Qy  721 TACGGCAAGCATGACATGAGGCCTTGCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  970 TACGGCAAGCATGACATGAGGCCTTGCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC  1029

Qy  781 AGACTGTACCATCCAGCCGCTGCACATCACAACACATTTACTTAACAGACAGTCACCGAC  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1030 AGACTGTACCCATCCAGCCGCTGCACATCACAACACATTTACTTAACAGACAGTCACCGAC  1089

Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCGGCCCCAGGCAAACTTGCACGACGCGC  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1090 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCGGCCCCAGGCAAACTTGCACGACGCGC  1149

Qy  901 TGCCAGGCGATTCGGGAGCCCCCTGGTGTCTGAACGATGGCCGCATGACTTGGTG  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1150 TGCCAGGCGATTCGGGAGCCCCCTGGTGTCTGAACGATGGCCGCATGACTTGGTG  1209

Qy  961 GGCATCATCAGCTGGGCTGGGCCTGGTGTGACAGAAGGATGTCCCGGGTGTGTACAAAG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1210 GGCATCATCAGCTGGGCTGGGCCTGGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAG 1269

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 1270 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1314
```

RESULT 9
E01937

LOCUS   E01937   1314 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01937
VERSION   E01937.1 GI:2170185
KEYWORDS   JP 198910416 7-A/4.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 4 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT   OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/4
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87   8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19),
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pTQkPAdeltatrp;
            FH   Key   Location/Qualifiers
            FH
            FT   source   1..1314
                                 /organism="synthetic construct"
            FT   CDS   1..1314
                                 /product='t-PA'.
FEATURES   Location/Qualifiers
     source   1..1314
                     /organism="synthetic construct"

Qy   961 GGCATCATCAGCTGGGCCTGGGCCTGGGCTGGGCTGGACGAAGGATGTCCCGGGTGTGTACACAAAG 1020
         |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db 1210 GGCATCATCAGCTGGGCCTGGGCCTGGGCTGGGCTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1269

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
         ||||||||||||||||||||||||||||||||||||||||||||
Db 1270 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1314

RESULT 10
LOCUS   A27725   1419 bp   DNA   linear   PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pmTQk112.
ACCESSION   A27725
VERSION   A27725.1 GI:2172236
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 49 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES   Location/Qualifiers
     source   1..1419
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   309 a   413 c   426 g   271 t
ORIGIN

Query Match   100.0%;   Score 1065;   DB 6;   Length 1419;
Best Local Similarity   100.0%;   Pred. No. 4.8e-248;
Matches 1065;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 355 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 414

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGTGCCTGGAATTCCATGATCCTGATAGGCAAGGTT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 415 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGTGCCTGGAATTCCATGATCCTGATAGGCAAGGTT 474

Qy 121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 475 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 534

Qy 181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 535 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 594

Qy 241 GAGTACTGTGATGTGCCGCTGCCTCCTGCTCGACTGCAGTACAGCAGCCTGCCAGCCTCAG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 595 GAGTACTGTGATGTGCCGCTGCCTCCTGCTCGACTGCAGTACAGCAGCCTGCCAGCCTCAG 654

Qy 301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCATC 360
        |||||||||||||||||||||||||||||||||||||||
Db 655 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCATC 714

Qy 361 TTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTCCGCACT 420
        |||||||||||||||||||||||||||||||||||||||||
Db 715 TTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTCCGCACT 774

Qy 421 TCCTGCTGGATTCTCTGCGCACTGGACTGCGAGTGTGAGCTCCGGC 480
        |||||||||||||||||||||||||||||||||||||||||||||
Db 775 TCCTGCTGGATTCTCTGCGCACTGGACTGCGAGTGTGAGCTCCGGC 834

Qy 481 ACGGTGATCTTGGGCAGAACATACCGGGGTGTCTGAACGATGCCGCATC 540
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 835 ACGGTGATCTTGGGCAGAACATACCGGGGTGTCTGAACGATGCCGCATC 894

```
Qy   181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG  594

Qy   241 GAGTACTGTGATGTGCCCTGTCCACTGCGCCTGAGACAGTACGCCAGCCTCAG  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 GAGTACTGTGATGTGCCCTGTCCACTGCGCCTGAGACAGTACGCCAGCCTCAG  654

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCAGCATCGCCTCCACCCCTGGCAGGCTGCCATC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   655 TTTCGCATCAAAGGAGGGCTCTTCGCCAGCATCGCCTCCACCCCTGGCAGGCTGCCATC  714

Qy   361 TTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGGCATCATCATCAGC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715 TTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCATCAGC  774

Qy   421 TCCTGCTGGATTCTCTCTGCCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   775 TCCTGCTGGATTCTCTCTGCCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTG  834

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGCAGAAATTTGAA  894

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATTGCG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATTGCG  954

Qy   601 CTGCTGCAGTCGAAATCGGATTCGTCCCGTGTGCCCGAGGAGCAGCGTGGTCGCACT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955 CTGCTGCAGTCGAAATCGGATTCGTCCCGTGTGCCCGAGGAGCAGCGTGGTCGCACT  1014

Qy   661 GTTGTGGCCTTCCCCCGGCCGGACCTGCAGCTGCCCGAGGAGTGTGAGCCTCCGGC  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1015 GTTGTGGCCTTCCCCCGGCCGGACCTGCAGCTGCCCGAGGAGTGTGAGCCTCCGGC  1074

Qy   721 TACGGCAAGCATCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTCATGTC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1075 TACGGCAAGCATCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTCATGTC  1134

Qy   781 AGACTGTACCCATCCAGCGCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1135 AGACTGTACCCATCCAGCGCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC  1194

Qy   841 AACATGCTGTGTGCTGGAGACACTCGGAGGCGGGGCCCCAGGCAAACTTGCACGACGCC  900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1195 AACATGCTGTGTGCTGGAGACACTCGGAGGCGGGGCCCCAGGCAAACTTGCACGACGCC  1254

Qy   901 TGCCAGGGGGATTCGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGAGTACTTGGTG  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255 TGCCAGGGGGATTCGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGAGTACTTGGTG  1314

Qy   961 GGCATCATCAGCTGGGCTGGGCTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315 GGCATCATCAGCTGGGCTGGGCTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1374

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA  1065
             |||||||||||||||||||||||||||||||||||||||||||||
Db  1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA  1419
```

```
Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGGG  954

Qy   601 CTGCTGCAGTCGAAATCGGATTCGTCCCGTGTGCCCGAGGAGCAGCGTGGTCCGCACT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955 CTGCTGCAGTCGAAATCGGATTCGTCCCGTGTGCCCGAGGAGCAGCGTGGTCCGCACT  1014

Qy   661 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCCGAGGAGTGTGAGCTTCCCGGC  720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1015 GTGTGCCTTCCCCCGGCCGGACCTGCAGCTGCCCGAGGAGTGTGAGCTTCCCGGC  1074

Qy   721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTCATGTC  780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1075 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGAGGCTCATGTC  1134

Qy   781 AGACTGTACCATCCAGCGCGCTGACATCACAACATTACTTAACAGAACAGTCACCGAC  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1135 AGACTGTACCATCCAGCGCGCTGACATCACAACATTACTTAACAGAACAGTCACCGAC  1194

Qy   841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGGGCCCCCAGGCAAACTTGCACGACGCC  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1195 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGGGCCCCCAGGCAAACTTGCACGACGCC  1254

Qy   901 TGCCAGGGCGATTCGGAGGCGCCCCCTGGTGTGTGTCTGACGATGGCCGCATGAGTACTTGGTG  960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255 TGCCAGGGGGATTCGGAGGCGCCCCCTGGTGTGTGTCTGACGATGGCCGCATGAGTACTTGGTG  1314

Qy   961 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1374

Qy  1021 GTTACCAACTACCTAGACTGATTCGTGACAACATGCCGACCGTGA  1065
             |||||||||||||||||||||||||||||||||||||||||||||
Db  1375 GTTACCAACTACCTAGACTGATTCGTGACAACATGCCGACCGTGA  1419
```

RESULT 11

AR059996
LOCUS   AR059996   1419 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 62 from patent US 5840533.
ACCESSION   AR059996
VERSION   AR059996.1 GI:598646
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1419)
  AUTHORS   Niwa.M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 62 24-NOV-1998;
FEATURES   Location/Qualifiers
     source   1..1419
                     /organism="unknown"
BASE COUNT   309 a   416 c   423 g   271 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1419;
Best Local Similarity   100.0%;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 12

E01944
LOCUS   E01944   1419 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01944
VERSION   E01944.1 GI:2170192
KEYWORDS   JP 198104167-A/11.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1419)

```
Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACGC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 TCTGAGGGAAACAGTGACTGCTACTTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC  414

Qy    61 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCGAATTCCATGATCCTGATAGGCAAGGTT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCTCGAATTCCATGATCCTGATAGGCAAGGTT  474

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCTGGGCAAACATAATTACTGCCGG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCTGGGCAAACATAATTACTGCCGG  534
```

```
AUTHORS     Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
TITLE       NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
JOURNAL     Patent: JP 1989104167-A 11 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT     OS  Artificial gene
            OC  Artificial sequence; Genes.
            PN  JP 1989104167-A/11
            PD  21-APR-1989
            PF  01-AUG-1988 JP 1988192320
            PR  03-AUG-1987 GB 87  8718298,  26-OCT-1987 GB 87  8725052; PR
            13-NOV-1987 GB 87  8726683
            PI  NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI  NOTANI JOJI,
            PI  KOBAYASHI MASAKAZU
            PC  C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19),
            PC  (C12N9/64,C12R1:91);
            CC  strandedness: Double;
            CC  topology: Linear;
            CC  hypothetical: No;
            CC  anti-sense: No;
            CC  *source: clone=pmTOk112;
            FH  Key             Location/Qualifiers
            FH
            FT  CDS             1..1419
            FT                  /product='new t-PA'.
FEATURES    source          1..1419
                            /organism="synthetic construct"
                            /mol_type="genomic DNA"
                            /db_xref="taxon:32630"
BASE COUNT     309 a    413 c    426 g    271 t
ORIGIN

Query Match        100.0%;  Score 1065;  DB 6;  Length 1419;
Best Local Similarity  100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    355 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 414

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCTGCCTCCCTCGGTGCCTCCTGCCTCCCTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    415 CTCACCGAGTCGGGTGCCTCCTGCCTCCCTGCCTCCCTCGGTGCCTCCTGCCTCCCTGCC 474

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    475 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCTGGGCCAAACATAATTACTGCCGG 534

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    535 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 594

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    595 GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 654

Qy    301 TTTCGGCATCAAAGGAGAGGGCTCTTCGCCGACATCGCCTGCCTCCACCCCTGCCGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655 TTTCGGCATCAAAGGAGAGGGCTCTTCGCCGACATCGCCTGCCTCCACCCCTGCCGCCATC 714

Qy    361 TTTGCCAAGCACAGGAGGTGCGCCGGAGAGTCGGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715 TTTGCCAAGCACAGGAGGTGCGCCGGAGAGTCGGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 774

Qy    421 TCCTGCTGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACACCTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 TCCTGCTGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACACCTG 834

Qy    481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    835 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTGAA 894

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACATTACGACAATGACATTGCG 954

Qy    601 CTGCTGCAGCTGAAATCGATCGTCCCGTGTCCCAGGAGAGCAGCGCTGGTCCGCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955 CTGCTGCAGCTGAAATCGATCGTCCCGTGTCCCAGGAGAGCAGCGCTGGTCCGCACT 1014

Qy    661 GTGTGGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGACGGAGTGTGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 GTGTGGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGACGGAGTGTGAGCTCTCCGGC 1074

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAACGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAACGGAGGCTCATGTC 1134

Qy    781 AGACTGTACCCATCAGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 AGACTGTACCCATCAGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1194

Qy    841 AACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 AACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1254

Qy    901 TGGCCAGGGCGAATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 TGGCCAGGGCGAATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGGACTTTGGTG 1314

Qy    961 GGCATCATCAGCTGGGCCTGGCTGTGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 GGCATCATCAGCTGGGCCTGGCTGTGACAGAGGATGTCCCGGGTGTGTACACAAAG 1374

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1419
```

```
RESULT 13
I06609
LOCUS       I06609                 1689 bp    DNA       linear    PAT 02-DEC-1994
DEFINITION  Sequence 44 from Patent EP 0293934.
ACCESSION   I06609
VERSION     I06609.1  GI:590639
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1689)
  AUTHORS   Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
            Hashimoto,A. and Yuzuriha,T.
  TITLE     Mutant t-PA with kringle replacement
  JOURNAL   Patent: EP 0293934-A1 44 07-DEC-1988;
FEATURES    Location/Qualifiers
     source          1..1689
                    /organism="unknown"
BASE COUNT     377 a    483 c    505 g    324 t
ORIGIN

Query Match        100.0%;  Score 1065;  DB 6;  Length 1689;
Best Local Similarity  100.0%;  Pred. No. 4.8e-248;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 684

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCTGCTTCCATGATCCTGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 CTCACCGAGTCGGGTGCCTCCTGCCTCCCTGCTTCCATGATCCTGATAGGCAAGGTT 744

Qy    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCTGGGCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

REFERENCE
AUTHORS Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
Hashimoto.A. and Yuzuriha.T.
TITLE Mutant t-PA with kringle replacement
JOURNAL Patent: EP 0293934-A1 46 07-DEC-1988;
FEATURES Location/Qualifiers
source 1..1689
/organism="unknown"

Unclassified.
1 (bases 1 to 1689)

BASE COUNT 377 a 484 c 504 g 324 t
ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1689;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGCTCAGCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGCTCAGCTACCGTGGCACGCACAGC 684

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 CTCACCGAGTCGGGTGCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 744

Qy    121 TACACAGCACAGAACCCAGTGCCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 TACACAGCACAGAACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG 804

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 864

Qy    241 GAGTACTGTGATGTGCCCTCTCCACCTGGGGCCTGAGACAGTACAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 GAGTACTGTGATGTGCCCTCTCCACCTGGGGCCTGAGACAGTACAGCCAGCCTCAG 924

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACACTGCCTCCACCCCTGCGAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    925 TTTCGCATCAAAGGAGGGCTCTTCGCCGACACTGCCTCCACCCCTGCGAGGCTGCCATC 984

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGCTCCGGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGCTCCGGGGGCATACTCATCAGC 1044

Qy    421 TCCTGCTGGATTCTCTCGCCGACTGCTTCGCCGACAGTTCCTCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TCCTGCTGGATTCTCTCGCCGACTGCTTCGCCGACAGTTCCTCAGC 1104

Qy    481 ACGGTGATCTTGGGCAGAACATACGGTGGTCCCTGCGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 ACGGTGATCTTGGGCAGAACATACGGTGGTCCCTGCGAGGAGCAGAAATTTGAA 1164

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1224

Qy    601 CTGCTGCAGCTGAAATGGCTGAAATCGGATTCGTGCCCAGGACTCGCTGCCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 CTGCTGCAGCTGAAATGGCTGAAATCGGATTCGTGCCCAGGACTCGCTGCCACT 1284

Qy    661 GTGTGCCTTCCCCGGGGACCTGCAGCTGCCGACTGGAACGCTGGACGAGTGTGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 GTGTGCCTTCCCCGGGGACCTGCAGCTGCCGACTGGAACGCTGGACGAGTGTGAGCTCTCCGGC 1344

Qy    721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTCATTCGGAGCGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTCATTCGGAGCGGCTCATGTC 1404

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGACAGTCACCGAC 1464

Qy    841 AACATGCTGTGTGCTGGAGACACTCGGACGCGCGGCCCCAGGCAAACTTGCGACGACGCC 900
          ||||||||||||||||
```

```
Db    745 TACACAGCACAGAACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG 804
          ||||||||||||||||
Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||
Db    805 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 864
          ||||||||||||||||
Qy    241 GAGTACTGTGATGTGCCCTCTCCACCTGGGGCCTGAGACAGTACAGCCAGCCTCAG 300
          ||||||||||||||||
Db    865 GAGTACTGTGATGTGCCCTCTCCACCTGGGGCCTGAGACAGTACAGCCAGCCTCAG 924
          ||||||||||||||||
Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACACTGCCTCCACCCCTGCGGGGGCTGTCCGGC 360
          ||||||||||||||||
Db    925 TTTCGCATCAAAGGAGGGCTCTTCGCCGACACTGCCTCCACCCCTGCGGGGGCTGTCCGGC 984
          ||||||||||||||||
Qy    361 CTGCTGCAGCTGAAATCGGATTCGTCCTCCCGCTGTGTCCGCAGCTGACCTGCAGCTGCCGACTGGAGCTCTCCGGC 420
          ||||||||||||||||
Db    985 CTGCTGCAGCTGAAATCGGATTCGTCCTCCCGCTGTGTCCGCAGCTGACCTGCAGCTGCCGACTGGAGCTCTCCGGC 1044
          ||||||||||||||||
Qy    721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAACGATGGCCATGACTTTGGTG 780
          ||||||||||||||||
Db   1345 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAACGATGGCCATGACTTTGGTG 1404
          ||||||||||||||||
Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGACAGTCACCGAC 840
          ||||||||||||||||
Db   1405 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGACAGTCACCGAC 1464
          ||||||||||||||||
Qy    961 GGCATCATCAGCTGGGGCTGTGGACAGAAGATGTCCCGGGGTGTGTACACAAAG 1020
          ||||||||||||||||
Db   1585 GGCATCATCAGCTGGGGCTGTGGACAGAAGATGTCCCGGGGTGTGTACACAAAG 1644
          ||||||||||||||||
Qy   1021 GTTACCAACTACCTAGACTGGATTCGTTGACAACATGCGACACCGTGA 1065
          ||||||||||||||||
Db   1645 GTTACCAACTACCTAGACTGGATTCGTTGACAACATGCGACACCGTGA 1689
```

RESULT 14
I06614
LOCUS I06614 1689 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 46 from Patent EP 0293934.
ACCESSION I06614
VERSION I06614.1 GI:590640
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```
Db   1465 AACATGCTGTGTGCTGGAGACACTCGGAGGCGGCGGCCCCAGGCAAACTTGACGACGCC 1524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1584

Qy    961 GGCATCATCAGCTGGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GGCATCATCAGCTGGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1644

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1645 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1689
```

RESULT 15
I08789
LOCUS   I08789   1689 bp   DNA   linear   PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent WO 8804690.
ACCESSION   I08789
VERSION   I08789.1 GI:588494
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1689)
  AUTHORS   Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.
  JOURNAL   Patent: WO 8804690-A 3 30-JUN-1988;
FEATURES   Location/Qualifiers
     source   1..1689
                     /organism="unknown"
BASE COUNT   379 a   480 c   505 g   325 t
ORIGIN

Query Match   100.0%; Score 1065; DB 6; Length 1689;
Best Local Similarity   100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTCAGCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTCAGCGTGGCACGCACAGC 684
```

Search completed: August 19, 2003, 04:46:33
Job time : 4473.78 secs

```
               GenCore version 5.1.6
      Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 375.192 Seconds
                                      (without alignments)
                           7662.485 Million cell updates/sec

Title: US-09-987-455-4
Perfect score: 1065
Sequence: 1 tctgaggaaacagtgactg.........gtgacaacatgcgaccgtga_1065

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_19Jun03:*
         1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
         2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
         3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
         4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
         5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
         6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
         7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
         8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
         9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
        10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
        11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
        12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
        13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
        14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
        15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
        16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
        17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
        18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
        19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
        20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
        21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
        22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
        23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
        24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
        25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1065	100.0	1065	24	AAD40616	Human kringle 2 se
2	1065	100.0	1068	10	AAN91120	Sequence of coding
3	1065	100.0	1170	10	AAN91135	Sequence of coding
4	1065	100.0	1314	10	AAN91123	Sequence of coding
5	1065	100.0	1419	10	AAN91133	Sequence of coding
6	1065	100.0	1680	17	AAT27588	Novel plasminogen
7	1065	100.0	1680	17	AAT27587	Novel plasminogen
8	1065	100.0	1689	9	AAN82178	Tissue plasminogen
9	1065	100.0	1780	11	AAQ01358	Sequence encoding
10	1065	100.0	2100	10	AAN91119	Sequence of native
11	1065	100.0	2100	11	AAQ05532	Plasmid pST12 enc
12	1065	100.0	2162	16	AAQ87370	Human tPA clone.
13	1065	100.0	2162	19	AAV37294	Human tissue plasm
14	1065	100.0	2547	7	AAN6O659	Sequence encoding
15	1065	100.0	2560	10	AAN90542	Plasmid pKG12 cont
16	1065	100.0	7533	8	AAN70491	Entire sequence of
17	1065	100.0	7533	9	AAN81970	Plasmid pEMp1-tPA.
18	1063.4	99.8	1614	11	AAQ03581	Glycosylated plasm
19	1063.4	99.8	1614	12	AAQ11550	Sequence encoding
20	1063.4	99.8	1659	12	AAQ11551	Sequence encoding
21	1063.4	99.8	1689	9	AAN82179	Tissue plasminogen
22	1063.4	99.8	1689	11	AAQ81090	Mutated t-PA analo
23	1063.4	99.8	1689	11	AAQ04903	Part of tPA024 gen
24	1063.4	99.8	1689	11	AAQ04904	Part of tPA023 gen
25	1063.4	99.8	1689	11	AAQ05177	Sequence encoding
26	1063.4	99.8	1689	12	AAQ13379	T-PA67+ mutant wit
27	1063.4	99.8	1689	24	ABK52546	Human tissue plasm
28	1063.4	99.8	1740	17	AAT27585	Full-length tissue
29	1063.4	99.8	1955	21	AAZ31998	Human tissue plasm
30	1063.4	99.8	1989	7	AAN60700	Sequence encoding
31	1063.4	99.8	2091	12	AAQ12073	T-PA variant contg
32	1063.4	99.8	2099	11	AAQ05534	Plasmid plgNTQTPA
33	1063.4	99.8	2100	12	AAQ12072	T-PA with -ve char
34	1063.4	99.8	2170	8	AAN70248	Sequence encoding
35	1063.4	99.8	2170	16	AAQ86576	Human tissue plasm
36	1063.4	99.8	2249	11	AAQ05535	Plasmid plgN delta
37	1063.4	99.8	2457	10	AAN60146	cDNA sequence enco
38	1063.4	99.8	2509	10	AAN90508	cDNA of human tiss
39	1063.4	99.8	2519	24	ABK52545	Human tissue plasm
40	1063.4	99.8	2519	24	ABL59544	Human tissue type
41	1063.4	99.8	2547	4	AAN30001	Sequence encoding
42	1063.4	99.8	7360	17	AAT15930	DHFR/intron (WTras
43	1062	99.7	1128	24	AAD40614	OmpA-K2S fusion pr
44	1062	99.7	1341	11	AAQ06762	Sequence encoding
45	1061.8	99.7	1068	10	AAN91122	Sequence of coding

ALIGNMENTS

```
RESULT 1
AAD40616
ID   AAD40616 standard; DNA; 1065 BP.
XX
AC   AAD40616;
XX
DT   30-OCT-2002  (first entry)
XX
DE   Human kringle 2 serine protease (K2S) DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; enzyme; human; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1065
FT                  /*tag= a
FT                  /product= "Human K2S protein"
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
```

```
Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660

Qy  661 GTGTGCCTTCCCCGGCGGCCGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GTGTGCCTTCCCCGGCGGCCGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 720

Qy  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGCTGAAGGAGGCTCATGTC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGCTGAAGGAGGCTCATGTC 780

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840

Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC 900

Qy  901 TGCCAGGGCGATTCGGGAGGCCCCCTGTGGTGTGTCTGAACGATGCCGATGACTTGGTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TGCCAGGGCGATTCGGGAGGCCCCCTGTGGTGTGTCTGAACGATGCCGATGACTTGGTG 960

Qy  961 GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA 1065
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA 1065
```

RESULT 2

AAN91120

ID AAN91120 standard; DNA; 1068 BP.
XX
AC AAN91120;
XX
DT 25-MAR-2003   (updated)
DT 03-OCT-2002   (updated)
DT 18-JUN-1990   (first entry)
XX
DE Sequence of coding region in plasmid pST112.
XX
KW Tissue plasminogen activator; tPA; thrombolytic agent;
KW plasminogen; vascular diseases.
XX
OS Synthetic.
XX
FH Key   Location/Qualifiers
FT CDS   1..1065
FT   /tag= a
XX
PN EP302456-A.
XX
PD 08-FEB-1989.
XX
PF 02-AUG-1988;   88EP-0112569.
XX
PR 03-AUG-1987;   87GB-0018298.
PR 26-OCT-1987;   87GB-0025052.
PR 13-NOV-1987;   87GB-0026683.
XX
PA (FUJI) FUJISAWA PHARM CO LTD.
XX
PI Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR WPI; 1989-040625/06.
DR P-PSDB; AAP94407.
XX
PT New tissue plasminogen activator -
PT comprising finger and growth factor domains lacking tPA for

---

XX
PI Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR WPI; 2002-519376/55.
DR P-PSDB; AAB25036.
XX
PT Producing active, correctly folded recombinant tissue plasminogen
PT activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT the protein-encoding DNA operably linked to DNA coding for signal
PT peptide OmpA -
XX
PS Claim 12; Page 31; 80pp; English.
XX
CC The present invention relates to a method of producing extracellularly
CC secreted, active, correctly folded, recombinant tissue plasminogen
CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC variants in prokaryotic cells by expressing the protein-encoding DNA
CC operably linked to DNA coding for signal peptide OmpA. The method is
CC useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC Sequences of the invention are useful for manufacturing a medicament
CC for treating stroke, cardiac infarction, acute myocardial infarction,
CC pulmonary embolism, any artery occlusion such as intracranial artery
CC occlusion (e.g., arteries supplying the brain), peripherally occluded
CC arteries, coronary artery occlusion, deep vein thrombosis or related
CC diseases associated with unwanted blood clotting. The present sequence
CC is human K2S DNA.
XX
SQ Sequence 1065 BP; 225 A; 314 C; 312 G; 214 T; 0 other;

Query Match   100.0%;   Score 1065;   DB 24;   Length 1065;
Best Local Similarity   100.0%;   Pred. No. 9.7e-237;
Matches 1065;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60

Qy  61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120

Qy 121 TACCAGCACAGAGAACCCCAGTGCCACTCGGGCCTGGGCCTGGCAAACATAATTACTGCCGG 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TACCAGCACAGAGAACCCCAGTGCCACTCGGGCCTGGGCCTGGCAAACATAATTACTGCCGG 180

Qy 181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGGCTGACGTGG 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGGCTGACGTGG 240

Qy 241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCTGAGACGAGTACAGCCAGCCTCAG 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCTGAGACGAGTACAGCCAGCCTCAG 300

Qy 301 TTTCGCATCAAAAGGAGGGCTCTTCGCGCCAGCATCATCGCCTCGCAGGCTGCCATC 360
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 TTTCGCATCAAAAGGAGGGCTCTTCGCGCCAGCATCATCGCCTCGCAGGCTGCCATC 360

Qy 361 TTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGGTTCCTCGTGCGGGGGCATACTCATCAGC 420
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGGTTCCTCGTGCGGGGGCATACTCATCAGC 420

Qy 421 TCCTGCTGGATTCTCTTCTCTGCCGCCCACTGCTTCCAGAGAGAGGTTTCCGCCCCACCTG 480
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 TCCTGCTGGATTCTCTTCTCTGCCGCCCACTGCTTCCAGAGAGAGGTTTCCGCCCCACCTG 480

Qy 481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCCGGAGGAGGCAGAAATTTGAA 540
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCCGGAGGAGGCAGAAATTTGAA 540

Qy 541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
```

```
Qy   901 TGCCAGGGCGGATTCGGAGGAGGCCCCCTGTGTGTCTGAACGATGCCGATGACTTGGTG 960
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   904 TGCCAGGGCGGATTCGGAGGAGGCCCCCTGTGTGTCTGAACGATGACGATGACTTGGTG 963

Qy   961 GGCATCATCAGCTGGGCGTGGGCCTGGGCTGTGGCACGAGAAGGATGTCCCGGGTGTACACAAAG 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964 GGCATCATCAGCTGGGCGTGGGCCTGGGCTGTGGCACGAGAAGGATGTCCCGGGTGTACACAAAG 1023

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1068
```

RESULT 3
AAN91135
ID   AAN91135 standard; DNA; 1170 BP.
XX
AC   AAN91135;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pmTrk.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..1065
FT   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988; 88EP-0112569.
XX
PR   03-AUG-1987; 87GB-0018298.
PR   26-OCT-1987; 87GB-0025052.
PR   13-NOV-1987; 87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94417.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1170 BP; 249 A; 338 C; 346 G; 237 T; 0 other;

Query Match   100.0%; Score 1065; DB 10; Length 1170;
Best Local Similarity   100.0%; Pred. No. 9.9e-237;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTGGCACGGCACGCACCAGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTGGCACGGCACGCACCAGC 165

Qy    61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 225
```

DT   03-OCT-2002   (updated)
DT   18-JUN-1990   (first entry)
XX
DE   Sequence of coding region in plasmid pTQkPA delta trp.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1065
FT   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88BP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94409.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page 7; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1314 BP; 286 A; 386 C; 393 G; 249 T; 0 other;

Query Match   100.0%;   Score 1065;   DB 10;   Length 1314;
Best Local Similarity   100.0%;   Pred. No. 1e-236;
Matches 1065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

[Sequence alignment data follows between Qy (query) and Db (database) rows, positions 1–1170.]

RESULT 4
AAN91123
ID   AAN91123 standard; DNA; 1314 BP.
XX
AC   AAN91123;
XX
DT   25-MAR-2003 (updated)

PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94416.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;

Query Match   100.0%;   Score 1065;   DB 10;   Length 1419;
Best Local Similarity   100.0%;   Pred. No. 1e-236;
Matches 1065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACGCAGC    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACGCAGC   414

Qy    61 CTCACCGAGTCGGGTGCCTCCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 CTCACCGAGTCGGGTGCCTCCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   474

Qy   121 TACAACAGCACAGAACCCCAGTGCCAGGCACTGGGCTCGGCAAACATAATTACTGCCGG   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 TACAACAGCACAGAACCCCAGTGCCAGGCACTGGGCTCGGCAAACATAATTACTGCCGG   534

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACACCGGAGGCTGACGTGG   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACACCGGAGGCTGACGTGG   594

Qy   241 GAGTACTGTGATGTGCCCCTGCTCCACCTGCGCGCCTGAGACAGTACAGCGCCAGCCTCAG   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 GAGTACTGTGATGTGCCCCTGCTCCACCTGCGCGCCTGAGACAGTACAGCGCCAGCCTCAG   654

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCCTCGCCGACATCGCCTGCAGGCTGCCATC   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   655 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCCTCGCCGACATCGCCTGCAGGCTGCCATC   714

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC   774

Qy   421 TCCTGCTGGATTCTCTCGCGCCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   670 TCCTGCTGGATTCTCTGCCGCGCCCCAGTGCTTCCTTCCAGGAGAGGTTTCCGCCCCACCACCTG   729

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   730 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA   789

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   790 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   849

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTCCCAGGAGAGCGCGTGGTCCGCACT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   850 CTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTCCCAGGAGAGCGCGTGGTCCGCACT   909

Qy   661 GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCAGCTCGACGAGTGTGAGCTCTCCGGC   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   910 GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCAGCTCGACGAGTGTGAGCTCTCCGGC   969

Qy   721 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   970 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC   1029

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1030 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   1089

Qy   841 AACATGCTGTGTGCTGGAGACACTCGGAGACGCGGCGGCCCCCAGGCAAACTTGCACGACGCC   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1090 AACATGCTGTGTGCTGGAGACACTCGGAGACGCGGCGGCCCCCAGGCAAACTTGCACGACGCC   1149

Qy   901 TGCCAGGGCGATTCGGGAGGCCCCCCTGTGTGTCTGAACGATGCCGCCATGACTTTGGTG   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1150 TGCCAGGGCGATTCGGGAGGCCCCCCTGTGTGTCTGAACGATGCCGCCATGACTTTGGTG   1209

Qy   961 GGCATCATCAGCTGGGGCCTGGGCTGCGTGTGACAGAAGGATGTCCCGGGTGTGTACAAAG   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1210 GGCATCATCAGCTGGGGCCTGGGCTGCGTGTGACAGAAGGATGTCCCGGGTGTGTACAAAG   1269

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA   1065
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1270 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA   1314
```

RESULT 5
AAN91133
ID   AAN91133 standard; DNA; 1419 BP.
XX
AC   AAN91133;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pmTQk112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1065
FT   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX

```
Qy    721 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1134

Qy    781 AGACTGTACCCATCCAGCCGCTGCAACATCACAACATTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 AGACTGTACCCATCCAGCCGCTGCAACATCACAACATTACTTAACAGAACAGTCACCGAC 1194

Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 AACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCGGCGCCCCAGGCAAACTTGCACGACGCC 1254

Qy    901 TGCCAGGGCGATTCGGAGGCCCCCTGGTGTGTCTGACAGGATGTGTCTGACAGGATGGCCGCATGACTTTGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 TGCCAGGGCGATTCGGAGGCCCCCTGGTGTGTCTGACAGGATGTGTCTGACAGGATGGCCGCATGACTTTGTG 1314

Qy    961 GGCATCATCAGCTGGGGCCTGGGCTGTGACAGAAGGATGTGCCCGGGTGTGTACAAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 GGCATCATCAGCTGGGGCCTGGGCTGTGACAGAAGGATGTGCCCGGGTGTGTACAAAAG 1374

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1419
```

RESULT 6
AAT27588
ID   AAT27588   standard; DNA; 1680 BP.
XX
AC   AAT27588;
XX
DT   25-MAR-2003 (updated)
DT   06-AUG-1996 (first entry)
XX
DE   Novel plasminogen activator DNA.
XX
KW   Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
KW   fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
KW   protein engineering; kringle; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   sig_peptide   1..105
FT   /*tag= a
FT   mat_peptide   106..1677
FT   /*tag= b
XX
PN   US5504001-A.
XX
PD   02-APR-1996.
XX
PF   06-JUN-1994; 94US-0254485.
XX
PR   25-NOV-1987; 87US-0125629.
PR   28-JAN-1992; 92US-0827587.
PR   06-JUN-1994; 94US-0254485.
XX
PA   (ZYMO) ZYMOGENETICS INC.
XX
PI   Foster DC;
XX
DR   WPI; 1996-187699/19.
DR   P-PSDB; AAR96223.
XX
PT   Hybrid plasminogen activator comprises human tPA activator and
PT   N-terminal crosslinking domain from alpha2-plasmin inhibitor -
PT   useful to treat thrombosis and image blood clots
XX
PS   Example 3; Fig 11; 35pp; English.
XX
CC   A DNA construct (AAT27586) codes for a novel plasminogen

CC   activator (AAR96222) in which the kringle K1 domain (AAR96221)
CC   of plasminogen replaces the native K1 domain of tissue
CC   plasminogen activator (tPA) (AAR96220). It was obtd. by
CC   inserting a synthetic plasminogen K1 domain DNA (AAT27586)
CC   into tPA cDNA and can be expressed in transformed Escherichia
CC   coli RRI cells (FERM P-9272). Novel plasminogen activators are
CC   produced that show increased clot lysing specificity or plasma
CC   half-life.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 1680 BP; 387 A; 471 C; 479 G; 343 T; 0 other;

Query Match   100.0%; Score 1065; DB 17; Length 1680;
Best Local Similarity   100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGC 675

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 735

Qy    121 TACAACAGCACAGAACCCCAGTGCCCAGGCACTCGGGCCTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 TACAACAGCACAGAACCCCAGTGCCCAGGCACTCGGGCCTGGGCCTGGGCAAACATAATTACTGCCGG 795

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCAGTGCTGTGAAGAACCGAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    796 AATCCTGATGGGATGCCAAGCCCTGGTGCCAGTGCTGTGAAGAACCGAGGCTGACGTGG 855

Qy    241 GAGTACTGCTGATGCCCTCCTGCTGCCACCTGCGGGCCTGAGACAGTAGCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    856 GAGTACTGCTGATGCCCTCCTGCTGCCACCTGCGGGCCTGAGACAGTAGCAGCCTCAG 915

Qy    301 TTTCGATCAAAGAGGAGGGCTCTTCGCCAGCATCGCCTCCACCCCTGGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    916 TTTCGATCAAAGAGGAGGGCTCTTCGCCAGCATCGCCTCCACCCCTGGCAGGCTGCCATC 975

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGAGAGGAGGTTCCTGTGCGGGGGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    976 TTTGCCAAGCACAGGAGGTCGCCCGAGAGGAGGTTCCTGTGCGGGGGCATACTCATCAGC 1035

Qy    421 TCCTGCTGGATTCTCTCGCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCACCACCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1036 TCCTGCTGGATTCTCTCGCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCACCACCTG 1095

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1096 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 1155

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATTGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1156 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATTGACATTGCG 1215

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTTCCCGGTGCTGTCCCAGGAGCAGCGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1216 CTGCTGCAGCTGAAATCGGATTCGTTCCCGGTGCTGTCCCAGGAGCAGCGGTCCGCACT 1275

Qy    661 GTGTGCCTTCCCCCGGACTGCAGCTGCCGAGCTGACGAGGTGTGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1276 GTGTGCCTTCCCCCGGACTGCAGCTGCCGAGCTGACGAGGTGTGAGCTCTCCGGC 1335

Qy    721 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1336 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1395

Qy    781 AGACTGTACCCATCCAGCCGCTGCAACATCACAACATTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1396 AGACTGTACCCATCCAGCCGCTGCAACATCACAACATTACTTAACAGAACAGTCACCGAC 1455

Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCC 900
```

SQ   Sequence 1680 BP; 386 A; 471 C; 480 G; 343 T; 0 other;

Query Match   100.0%;   Score 1065; DB 17; Length 1680;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTGGAATGGTCAGCCTACCGTGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 TCTGAGGGAAACAGTGACTGCTACTTGGAATGGTCAGCCTACCGTGCACGCACAGC 675

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCCGTGGAATTCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676 CTCACCGAGTCGGGTGCCTCCTGCCCGTGGAATTCATGATCCTGATAGGCAAGGTT 735

Qy    121 TACACAGCACAGACCCCAGTGCCCAGGCACTGGCCTGGGCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 TACACAGCACAGACCCCAGTGCCCAGGCACTGGCCTGGGCAAACATAATTACTGCCGG 795

Qy    181 AATCCTGATGGGATGCCAAGCCCTGGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    796 AATCCTGATGGGATGCCAAGCCCTGGTGCTGAAGAACCGCAGGCTGACGTGG 855

Qy    241 GAGTACTGTGATGTGCCCTGCTTCCACCTGCGCGCCTGAGACAGTACAGCCAGCCTCAG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    856 GAGTACTGTGATGTGCCCTGCTTCCACCTGCGCGCCTGAGACAGTACAGCCAGCCTCAG 915

Qy    301 TTTCGCATCAAAGAGGGCTCTTGCGCAGCATGCTGGCTCCACCCCTGGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    916 TTTCGCATCAAAGAGGGCTCTTGCGCAGCATGCTGGCTCCACCCCTGGCAGGCTGCCATC 975

Qy    361 TTTGCCAAGCACAAGAGGTCGCCGCGCGGTTCCTGTGCGGGGGCCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    976 TTTGCCAAGCACAAGAGGTCGCCGCGCGGTTCCTGTGCGGGGGCCATACTCATCAGC 1035

Qy    421 TCCTGCTGGATTCTCTCGCGCAGAGAGGTTTCCGCCCCACCACCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1036 TCCTGCTGGATTCTCTCGCGCAGAGAGGTTTCCGCCCCACCACCTG 1095

Qy    481 ACGGTGATCTTGGGCAGACATACGGTGGTCCTGGCGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1096 ACGGTGATCTTGGGCAGACATACGGTGGTCCTGGCGAGGAGCAGAAATTTGAA 1155

Qy    541 GTGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1156 GTGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1215

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCGGTGGTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1216 CTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCGGTGGTCCGCACT 1275

Qy    661 GTGTGCCTTCCCCGGCCGACTGCAGCTGACGGAGTGTGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1276 GTGTGCCTTCCCCGGCCGACTGCAGCTGACGGAGTGTGAGCTCTCCGGC 1335

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1336 TACGGCAAGCATGAGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 1395

Qy    781 AGACTGTACCATCCAGCCGCTGACATCACAACATTACTTAACAGAACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1396 AGACTGTACCATCCAGCCGCTGACATCACAACATTACTTAACAGAACAGTCACCGAC 1455

Qy    841 AACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGAGCGC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1456 AACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGAGCGC 1515

Qy    901 TGCCAGGGGATTCGGAGCGTGTCTGAACGGCCGCATGACGTTTGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1516 TGCCAGGGGATTCGGAGCGTGTCTGAACGGCCGCATGACGTTTGGTG 1575

Qy    961 GGCATCATCAGCTGGGGCCTGGCGTGTGACAAGGATGTCCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1576 GGCATCATCAGCTGGGGCCTGGCGTGTGACAAGGATGTCCCGGGTGTGTACACAAAG 1635

Db   1456 AACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGAGCGC 1515

Qy    901 TGCCAGGGGATTCGGAGCGTGTCTGACAGAAGGATGTCCGGGTGTGTACACAAAG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1576 TGCCAGGGGATTCGGAGCGTGTCTGACAGAAGGATGTCCGGGTGTGTACACAAAG 1635

Qy    961 GGCATCATCAGCTGGGGCCTGGCGTGTGACAAACATGCCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1636 GGCATCATCAGCTGGGGCCTGGCGTGTGACAAACATGCCGACCGTGA 1680
```

RESULT 7
AAT27587
ID   AAT27587 standard; DNA; 1680 BP.
XX
AC   AAT27587;
XX
DT   25-MAR-2003 (updated)
DT   06-AUG-1996 (first entry)
XX
DE   Novel plasminogen activator DNA.
XX
KW   Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
KW   fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
KW   protein engineering; kringle; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   sig_peptide   1..105
FT   /*tag= a
FT   mat_peptide   106..1677
FT   /*tag= b
XX
PN   US5504001-A.
XX
PD   02-APR-1996.
XX
PF   06-JUN-1994; 94US-0254485.
XX
PR   25-NOV-1987; 87US-0125629.
PR   28-JAN-1992; 92US-0827587.
PR   06-JUN-1994; 94US-0254485.
XX
PA   (ZYMO) ZYMOGENETICS INC.
XX
PI   Foster DC;
XX
DR   WPI; 1996-187699/19.
DR   P-PSDB; AAR96222.
XX
PT   Hybrid plasminogen activator comprises human tPA activator and
PT   N-terminal crosslinking domain from alpha2-plasmin inhibitor -
PT   useful to treat thrombosis and image blood clots
XX
PS   Example 3; Fig 10; 35pp; English.
XX
CC   A DNA construct (AAT27587) codes for a novel plasminogen
CC   activator (AAR96223) in which the kringle K1 domain of plasminogen
CC   (see also AAR96221) is mutated to substitute asparagine for aspartic
CC   acid at position 5, and replaces the native K1 domain of tissue
CC   plasminogen activator (tPA) (AAR96220). It was obtd. by mutagenesis
CC   of the sequence in vector Zem99-8000 (see also AAT27587). The novel
CC   protein can be expressed in Escherichia coli RRI/Zem99-8010 cells
CC   (FERM P-9315). Novel plasminogen activators are produced that show
CC   increased clot lysing specificity or plasma half-life.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX

Db   745 TACAGACAGAACCCCAGTGCCCAGGCACTGGCCTGGCAAACATAATTACTGCGGG 804
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 AATCCTGATGGGATGCCAAGCCTGGTTGCCAGTGCTGTCGAAGAACCGAGGCTGACGTGG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 AATCCTGATGGGATGCCAAGCCCTGGTTGCCAGTGCTGTGAAGAACCGAGGCTGACGTGG 864

Qy   241 GAGTACTGTGATGTGCCCTGCTCCACCTGCGGCCTGAGACAGCTACAGCCAGCCTCAG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 GAGTACTGTGATGTGCCCTGCTCCACCTGCGGCCTGAGACAGCTACAGCCAGCCTCAG   924

Qy   301 TTTCGCATCAAAGGAGGGCTCTTGCCACGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 TTTCGCATCAAAGGAGGGCTCTTCGCCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 984

Qy   361 TTTGCCAAGCACAGGAGTGCGCCGGACGGTTCCTGCGTGTGCGGGGGCATACTCATCAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985 TTTGCCAAGCACAGGAGTGCGCCGGAGCGGTTCCTGCGTGTGCGGGGGCATACTCATCAGC 1044

Qy   421 TCCTGCTGGATTCTCTTCGCGCCCACTGCTTCCAGGAGAGGTTCCGCCCACCACCTG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TCCTGCTGGATTCTCTTCGCGCCCACTGCTTCCAGGAGAGGTTCCGCCCCACCACCTG   1104

Qy   481 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 1164

Qy   541 GTCGAAAAATACATTGTCCATAAGGAGAATTCGATTCGCCCGTGTGCCCAGAGAGCAGCGCTGGTCCGCACT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 GTCGAAAAATACATTGTCCATAAGGAGAATTCGATTCGCCCGTGTGCCCAGAGAGCAGCGCTGGTCCGCACT 1224

Qy   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGCCGACTGACGTGGTCGTCCGCACT   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 CTGCTGCAGCTGAAATCGGATTCGTCCCGGCCGACTGACGTGGTCGTCCGCACT 1284

Qy   661 GTGTGCCTTCCCCGGGACTGCAGCTGCAGCTGCAGCTGACGAGTGTGACGGAGTGTGAGCTCTCCGGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 GTGTGCCTTCCCCGGGACTGCAGCTGCAGCTGCAGCTGACGAGTGTGACGGAGTGTGAGCTCTCCGGC 1344

Qy   721 TACGGCAAGCAATGCATGAGGAGCCCTTGCTCTCCTTTCTATTCGGAGCGCGCTGAAGGAGGCTCATGTC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 TACGGCAAGCAATGCATGAGGAGCCCTTGCTCTCCTTTCTATTCGGAGCGCGCTGAAGGAGGCTCATGTC 1404

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 AGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1464

Qy   841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGGCC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1524

Qy   901 TGCCAGGCGGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGCCATGACTTTGGTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 TGCCAGGCGGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGCCATGACTTTGGTG 1584

Qy   961 GGCATCATCAGCTGGGGCTGGGGCTGTGACAGAGGATGTCCCGGGTGTGTACAACAAAG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GGCATCATCAGCTGGGGCTGGGGCTGTGACAGAGGATGTCCCGGGTGTGTACAACAAAG 1644

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1645 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1689

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
           |||||||||||||||||||||||||||||||||||||||||||
Db   1636 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1680

RESULT 8
AAN82178
ID   AAN82178 standard; DNA; 1689 BP.
XX
AC   AAN82178;
XX
DT   25-MAR-2003 (updated)
DT   10-MAR-2003 (updated)
DT   03-NOV-1990 (first entry)
XX
DE   Tissue plasminogen activator with S-119 substd for M.
XX
KW   Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1689
FT   /*tag= a
FT   /product=modified tPA
XX
PN   JP63230083-A.
XX
PD   26-SEP-1988.
XX
PF   20-MAR-1987;   87JP-0064339.
XX
PR   20-MAR-1987;   87JP-0064339.
XX
PA   (EISA) EISAI CO LTD.
XX
DR   WPI; 1988-311961/44.
DR   P-PSDB; AAP82581.
XX
PT   Modified tissue plasminogen activator - having glycine-183 and serine-186
PT   residues sustd. with serine and threonine.
XX
PS   Disclosure; Page ?; 16pp; Japanese.
XX
CC   One N-glycosylation site, i.e. NSS (117-119) is substituted with
CC   NSM and the N-glycosylation is removed.
CC   Plasmid encoding the modified tPA is 99-6300 and its transformant is
CC   E.coli RR1-Zem 99-6300 (FERM P-9127).
CC   This modified tPA, used to treat thrombosis, is of high quality and
CC   has a longer half life period in blood.
CC   See also AAN82177-N82179.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

Query Match   100.0%; Score 1065; DB 9; Length 1689;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTTCAGCCTGGCACGCACAGC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTTCAGCCTGGCACGCACAGC 684

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCGCCTCCCGTGAATCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685 CTCACCGAGTCGGGTGCCTCCTGCGCCTCCCGTGAATCCATGATCCTGATAGGCAAGGTT 744

Qy   121 TACAGACAGCAGAACCCCAGTGCCCAGGCACTGGCCTGGCAAACATAATTACTGCCGG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9
AAQ01358
ID   AAQ01358 standard; DNA; 1780 BP.
XX
AC   AAQ01358;
XX
DT   25-MAR-2003 (updated)
DT   08-SEP-1990 (first entry)
XX
DE   Sequence encoding wild type tissue plasminogen activator. (t-PA).

```
XX
KW   Wild type tissue plasminogen activator (t-PA); infarction treatment;
KW   thrombosis treatment; embolism treatment.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   misc_feature   1..123
FT                  /*tag= a
FT                  /note="synthetic 5' adaptor"
FT   misc_feature   1738..1780
FT                  /*tag= b
FT                  /note="synthetic 3' adaptor"
XX
PN   EP351246-A.
XX
PD   17-JAN-1990.
XX
PF   14-JUL-1989;   89EP-0307194.
XX
PR   15-JUL-1988;   88DK-0003952.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Petersen LC, Boel E;
XX
DR   WPI; 1990-016567/03.
DR   P-PSDB; AAR04699.
XX
PT   New tissue plasminogen activator (t-PA) analogue -
PT   with higher fibrin selectivity than native t-pa, useful for
PT   treating infarction, thrombosis and embolism
XX
PS   Disclosure; Fig 5A-D; 24pp; English.
XX
CC   It is modified in the patent by replacing one or more codons specifying
CC   a positively charged amino acid.  The resulting analogues are inserted
CC   into a replicable expression vector which is used to transform or
CC   transfect a host cell which is grown to express a t-PA analogue.  The
CC   t-PA analogue is useful for the treatment of diseases or disorders
CC   associated with the formation of thrombi in blood vessels e.g.
CC   infarctions, thrombosis and embolism.  The analogue, in the 1-chain form,
CC   exhibits the properties of a proenzyme; however on plasmin-catalysed
CC   cleavage of the one-chain form, the activity of the 2-chain form is fully
CC   retained.  Compared to native t-PA, it has a higher fibrin selectivity,
CC   this results in a fibrinolytic agent with a higher fibrin selectivity,
CC   as the fibrinogenolytic activity induced by the analogue is reduced
CC   relative to that induced by native 1-chain t-PA.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;

Query Match       100.0%;  Score 1065;  DB 11;  Length 1780;
Best Local Similarity  100.0%;  Pred. No. 1.1e-236;
Matches 1065;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    641 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 700

Qy     61 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    701 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 760

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    761 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 820

Qy    181 AATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    821 AATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 880

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
```

```
Db    881 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTCGCATCAAAGAGGGGCTCTTGCCCGAGACGGTTCCTGTGGGGGCATACTCATCAGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941 TTTCGCATCAAAGAGGGGCTCTTGCCCGAGACGGTTCCTGTGGGGGCATACTCATCAGC 1000

Qy    361 TTTGCCAAGCACAGAGGAGTCGCCCGAGAGCGGTTCCTGTGTGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1001 TTTGCCAAGCACAGAGGAGTCGCCCGAGAGCGGTTCCTGTGTGGGGCATACTCATCAGC 1060

Qy    421 TCCTGCTGGATTCTTCTCTGCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACCTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1061 TCCTGCTGGATTCTTCTCTGCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACCTG 1120

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGGCGGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1121 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGGCGGAGGAGCAGAAATTTGAA 1180

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1181 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACAATGACATTGCG 1240

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTTCCCGTGTGCCGAGAGCAGCGTGTCGCCGCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1241 CTGCTGCAGCTGAAATCGGATTCGTTCCCGTGTGCCGAGAGCAGCGTGTCGCCGCACT 1300

Qy    661 GTGTGCCTTCCCCGGCGGACCTGCCAGCTGCGACGCTGCCGGACTGGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1301 GTGTGCCTTCCCCGGCGGACCTGCCAGCTGCGACGCTGCCGGACTGGAGCTCTCCGGC 1360

Qy    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1361 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1420

Qy    781 AGACTGTACCCATCACGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1421 AGACTGTACCCATCACGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1480

Qy    841 AACATGCTGTGTCGAGACACTCGGAGACACTTGCACGACGCGCCAAACTTGCACGACGCC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1481 AACATGCTGTGTCGAGACACTCGGAGACACTTGCACGACGCGCCAAACTTGCACGACGCC 1540

Qy    901 TGGCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCGAGAAGCATGGCCGATGGACTTTGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1541 TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCGAGAAGCATGGCCGATGGACTTTGGTG 1600

Qy    961 GGCATCATCAGCTGGGGCCCTGGGGCTGTGGACAGAGAAGGATGTCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1601 GGCATCATCAGCTGGGGCCCTGGGGCTGTGGACAGAGAAGGATGTCCGGGTGTGTACACAAAG 1660

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTCGTGACAACATGCGACACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1661 GTTACCAACTACCTAGACTGGATTCGTCGTGACAACATGCGACACCGTGA 1705
```

```
RESULT 10
AAN91119
ID   AAN91119 standard; DNA; 2100 BP.
XX
AC   AAN91119;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of native tPA in plasmid pSTI12.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
```

```
Qy    541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1248

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCGTGGTCCGCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1249 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCGTGGTCCGCACT 1308

Qy    661 GTGTGGCCTTCCCCCGGGGGACCTGCAGCTGCCGCGACTGAGAGTGTGAGCTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1309 GTGTGGCCTTCCCCCGGGGGACCTGCAGCTGCCGCGACTGAGAGTGTGAGCTCTCCGGC 1368

Qy    721 TACGGCAAGCATGAGGGCCTTGCTCCTTTCTATTCGGAGCGGCGCTGAAGGAGGGCTCATGTC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369 TACGGCAAGCATGAGGGCCTTGCTCCTTTCTATTCGGAGCGGCGCTGAAGGAGGGCTCATGTC 1428

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1488

Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCAGGCAAACTTGCACGACGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCAGGCAAACTTGCACGACGCC 1548

Qy    901 TGCCAGGGGGATTGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549 TGCCAGGGGGATTCGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1608

Qy    961 GGCATCATCAGCTGGGGCCTGTGGCTGTGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1609 GGCATCATCAGCTGGGGCCTGTGGCTGTGACAGAGGATGTCCCGGGTGTGTACACAAAG 1668

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1713
```

```
RESULT 11
AAQ05532
ID   AAQ05532 standard; DNA; 2100 BP.
XX
AC   AAQ05532;
XX
DT   25-MAR-2003  (updated)
DT   11-DEC-1990  (first entry)
XX
DE   Plasmid pST112 encoding novel N-terminal for tissue plasminogen
DE   activator (tPA).
XX
KW   Fibrin; clotting; thrombolytic; vascular disease; stroke;
KW   myocardial infarction; heart attack; pulmonary embolism; ds;
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             25..1710
FT                   /*tag= a
FT   mat_peptide     130..1710
FT                   /*tag= b
XX
PN   EP379890-A.
XX
PD   01-AUG-1990.
XX
PF   10-JAN-1990;    90EP-0100457.
XX
PR   23-JAN-1989;    89GB-0001422.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Satoh S, Suzuki S, Otsuka K, Kusunoki C;
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             25..1710
FT                   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;    88EP-0112569.
XX
PR   03-AUG-1987;    87GB-0018298.
PR   26-OCT-1987;    87GB-0025052.
PR   13-NOV-1987;    87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94406.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Example 29; Fig 21; 68pp; English.
XX
CC   cDNA sequence of native tPA gene is excised from plasmid pST112, and
CC   digested with BamHI and SalI to form plasmid pST118.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;
```

```
Query Match             100.0%;  Score 1065;  DB 10;  Length 2100;
Best Local Similarity   100.0%;  Pred. No. 1.1e-236;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGGCACGCACAGC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGGCACGCACAGC 708

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCTCCTGCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 CTCACCGAGTCGGGTGCCTCCTGCTCCTGCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 768

Qy    121 TACAGCAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 TACAGCAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 828

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAAGAACCGCAGGCTGACGTGG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAAGAACCGCAGGCTGACGTGG 888

Qy    241 GAGTACTGTGATGTGCCCCTCTGCTCCACAGTACAGCAGACAGCAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 GAGTACTGTGATGTGCCCCTCTGCTCCACAGTACAGCAGACAGCAGCCAGCCTCAG 948

Qy    301 TTTCGCATCAAAGAGGAGGGGCTCTTCGCCGACATGCCTCGCACCCCTGCCAGGCTGCCATC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    949 TTTCGCATCAAAGAGGAGGGGCTCTTCGCCGACATGCCTCGCACCCCTGCCAGGCTGCCATC 1008

Qy    361 TTTGCCAAGCACAGGAGTGCGCCCGGACGGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009 TTTGCCAAGCACAGGAGTGCGCCCGGACGGCCGGTTCCTGTGCGGGGCATACTCATCAGC 1068

Qy    421 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCGCCGGGGTTTCCGCCCCACCCTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1069 TCCTGCTGGATTCTCTCTGCGCCCACTGCTTCGCCGGGGTTTCCGCCCCACCCTG 1128

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGCAGGGAGGAGCAGAAATTTGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGCAGGGAGGAGCAGAAATTTGAA 1188
```

DR　　WPI; 1990-232757/31.
DR　　P-PSDB; AAR06237.
XX
PT　　New tissue plasminogen activator - having N-terminal peptide of
PT　　plasminogen linked to tissue plasminogen activator for increased
PT　　stability in vivo.
XX
PS　　Disclosure; Page 7; 7pp; English.
XX
CC　　tPA with N-terminal peptide of plasminogen linked is more stable in
CC　　vivo than the native form. It is useful as a thrombolytic agent in
CC　　the treatment of vascular diseases eg myocardial infarction,
CC　　pulmonary embolism etc.
CC　　(Updated on 25-MAR-2003 to correct PA field.)
XX
SQ　　Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match　　　100.0%;　Score 1065;　DB 11;　Length 2100;
Best Local Similarity　100.0%;　Pred. No. 1.1e-236;
Matches 1065;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　　1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　649 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 708

Qy　　 61 CTCACCGAGTCGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　709 CTCACCGAGTCGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 768

Qy　　121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　769 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 828

Qy　　181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACCTGG 240
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　829 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACCTGG 888

Qy　　241 GAGTACTGTGATGTGCCCTCGTCTCCACCTGGGCCTCGAGACAGTACAGCCAGCCTCAG 300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　889 GAGTACTGTGATGTGCCCTCGTCTCCACCTGGGCCTCGAGACAGTACAGCCAGCCTCAG 948

Qy　　301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGGCCTCCACCCCTGCAGGCTGCCATC 360
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　949 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGGCCTCCACCCCTGCAGGCTGCCATC 1008

Qy　　361 TTTGCCAAGACAGGAGGTCGGCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGC 420
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1009 TTTGCCAAGACAGGAGGTCGGCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGC 1068

Qy　　421 TCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCCGGTCCTGGAGGAGGTTTCCGCCCACCTG 480
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1069 TCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCCGGTCCTGGAGGAGGTTTCCGCCCACCTG 1128

Qy　　481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 540
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1129 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 1188

Qy　　541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1189 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1248

Qy　　601 CTGCTGCAGCTCGAAATCGGATTCGTCCCGGAGAGCAGCGCTGGTCCGGCACT 660
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1249 CTGCTGCGAGCTCGAAATCGGATTCGTCCCGGAGAGCAGCGCTGGTCGGCACT 1308

Qy　　661 GTGTGCCTTCCCCCGGCGACCTGCCAGCTGCCGACTGACGAGTGTGAGCTCTCCGGC 720
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1309 GTGTGCCTTCCCCCGGCGACCTGCCAGCTGCCGACTGACGAGTGTGAGCTCTCCGGC 1368

Qy　　721 TACGGCAAGCAATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC 780
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1369 TACGGCAAGCAATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC 1428

Qy　　781 AGACTGTACCCATCCAGCGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 840
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1429 AGACTGTACCCATCCAGCGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC 1488

Qy　　841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGGCGGCGGCGGCGGCAAACTTGCACGACGCC 900
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1489 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGGCGGCGGCGGCGGCAAACTTGCACGACGCC 1548

Qy　　901 TGCCAGGGGGATTGGGGAGGCCCCCTGGTGTCTGAACGATGGCCGATGACTTGGTG 960
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1549 TGCCAGGGGGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGATGACTTTGGTG 1608

Qy　　961 GGCATCATCAGCTGGGGCTGGGCTGGCTGTGACAGAAGGATGTCCCGGGTGTGTACCACAAG 1020
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1609 GGCATCATCAGCTGGGGCTGGGCCTGGCTGTGACAGAAGGATGTCCCGGGTGTGTACCACAAG 1668

Qy　1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
　　　　　|||||||||||||||||||||||||||||||||||||||||||||
Db　1669 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1713

RESULT 12
AAQ87370
ID　　AAQ87370 standard; DNA; 2162 BP.
XX
AC　　AAQ87370;
XX
DT　　25-MAR-2003　(updated)
DT　　19-SEP-1995　(first entry)
XX
DE　　Human tPA clone.
XX
KW　　Intron; recombination; combinatorial gene; trans-splicing;
KW　　gene therapy; polymerase chain reaction; PCR; primer; amplification;
KW　　tissue plasminogen activator; tPA; plasmid TPA-KS+; thrombolytic;
KW　　ds.
XX
OS　　Synthetic.
XX
FH　　Key　　　　　　Location/Qualifiers
FT　　sig_peptide　82..334
FT　　　　　　　　/*tag= a
FT　　　　　　　　/note="signal sequence and finger-like domain"
FT　　　　　　　　335..447
FT　　　　　　　　/*tag= b
FT　　CDS　　　　　/product= EGF-like domain
FT　　　　　　　　448..714
FT　　　　　　　　/*tag= c
FT　　CDS　　　　　/product= Kringle-1 domain
FT　　　　　　　　715..972
FT　　　　　　　　/*tag= d
FT　　CDS　　　　　/product= Kringle-2 domain
FT　　　　　　　　973..2162
FT　　　　　　　　/*tag= e
FT　　CDS　　　　　/product= catalytic domain
XX
PN　　WO9507351-A1.
XX
PD　　16-MAR-1995.
XX
PF　　12-SEP-1994; 94WO-US10146.
XX
PR　　10-SEP-1993; 93US-0119512.
XX
PA　　(HARD) HARVARD COLLEGE.
XX
PI　　Jarrell KA;
XX
DR　　WPI; 1995-123425/16.
XX
PT　　New intron-mediated recombinant techniques - used for the
PT　　generation and selection of novel genes and gene prods. for use
PT　　in therapy

```
XX
PS   Example 4; Page 59-60; 87pp; English.
XX
CC   A cDNA clone of human tissue plasminogen activator (tPA) was
CC   amplified by PCR using the primers given in AAQ87368-69. The
CC   amplified tPA DNA (AAQ87370) was ligated into vector KS+ to
CC   obtain plasmid TPA-KS+. The construct was used in combinatorial
CC   methods involving RNA splicing-mediated shuffling of tPA domains
CC   in plasmid pINVI (AAQ87347) to generate novel tPAs having
CC   improved thrombolytic properties.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match             100.0%;  Score 1065;  DB 16;  Length 2162;
Best Local Similarity   100.0%;  Pred. No. 1.1e-236;
Matches 1065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACGTGGCACGCACAGC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACGTGGCACGCACAGC 765

Qy     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT 825

Qy    121 TACCACGACCAGAACCCCAGTGCCCAGCACTCGGCCTGGGCCAAACATAATTACTGCGGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TACCACGACCAGAACCCCAGTGCCCAGCACTCGGCCTGGGCCAAACATAATTACTGCGGG 885

Qy    181 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGGTGACGTGG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGGTGACGTGG 945

Qy    241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 1005

Qy    301 TTTCGGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGCTGCCATC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TTTCGGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGCTGCCATC 1065

Qy    361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCATACTCATCAGC 1125

Qy    421 TCCTGCTGGATTCTCTCTGCCCGCCACTGCTTCTCCAGGAGAGGTTTCCGCCCCACCACTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 TCCTGCTGGATTCTCTCTGCCCGCCACTGCTTCTCCAGGAGAGGTTTCCGCCCCACCACTG 1185

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 1245

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1305

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 1365

Qy    661 GTGTGCCTTCCCCCGGACCTGCAGCTGACCGCCGGACTGGACGAGTGTGAGCGTCTCCGGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 GTGTGCCTTCCCCCGGACCTGCAGCTGACCGCCGGACTGGACGAGTGTGAGCGTCTCCGGC 1425

Qy    721 TACGGCAAGCATGAGGGCCTTGTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 TACGGCAAGCATGAGGGCCTTGTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 1485

Qy    781 AGACTGTACCCATCCAGCCGCGCTGCCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 AGACTGTACCCATCCAGCCGCGCTGCCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1545
```

```
Qy    841 AACATGTCGTGTGTCTGAGAGACACTCGGACGCGGGGGCGGCCCAGGCAAACTTGCACGACGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 AACATGTCGTGTGTCTGGGAGACACTCGGACGCGGGGGCGGCCCAGGCAAACTTGCACGACGCC 1605

Qy    901 TGCCAGGGCGGATTCGGGAGGCCCCCTGGTGTCTGAACGACGATGGCCGATGACTTTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 TGCCAGGGCGGATTCGGGAGGCCCCCTGGTGTCTGAACGACGATGGCCGATGACTTTGGTG 1665

Qy    961 GGCATCATCAGCTGGGCGGCCTGGGCTGTGGACAGAGGATGTCCCGGGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 GGCATCATCAGCTGGGCGGCCTGGGCTGTGGACAGAGGATGTCCCGGGTGTACACAAAG 1725

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACACATGCGACCGGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGGTGA 1770
```

```
RESULT 13
AAV37294
ID   AAV37294 standard; DNA; 2162 BP.
XX
AC   AAV37294;
XX
DT   10-SEP-1998 (first entry)
XX
DE   Human tissue plasminogen activator gene sequence.
XX
KW   Plasmid pINVI; reverse-splicing intron; group II intron;
KW   exon binding site; domain V motif; branch site acceptor;
KW   nucleophilic group; transesterification; phosphodiester bond;
KW   autocatalytic Y-branched intron; reverse splicing reaction; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    82..334
FT                   /*tag= a
FT                   /note= "signal sequence and finger-like domain"
FT   misc_feature    335..447
FT                   /*tag= b
FT                   /note= "EGF-like domain"
FT   misc_feature    448..714
FT                   /*tag= c
FT                   /note= "Kringle-1 domain "
FT   misc_feature    715..972
FT                   /*tag= d
FT                   /note= "Kringle-2 domain"
FT   misc_feature    973..2162
FT                   /*tag= e
FT                   /note= "catalytic domain"
XX
PN   US5780272-A.
XX
PD   14-JUL-1998.
XX
PF   07-JUN-1995; 95US-0488015.
XX
PR   10-SEP-1993; 93US-0119512.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Jarrell KA;
XX
DR   WPI; 1998-413060/35.
XX
PT   Reverse splicing construct containing fragments of autocatalytic
PT   introns - able to cleave and ligate discontinuous nucleic acid for
PT   generating new genes and e.g. ribozymes, libraries of enzymes and
PT   antibodies
XX
PS   Example 4; Columns 53-56; 56pp; English.
XX
```

```
CC   The present sequence represents the human tissue plasminogen activator
CC   gene. It was used to construct plamsid TPA-KS+, which is used in the
CC   course of the invention. The specification describes a purified
CC   reverse-splicing intron which comprises a segment comprising a
CC   5'-part of a group II intron, including an exon binding site not
CC   naturally present in the intron and a second segment comprising a
CC   3'-part of a group II intron, including a domain V motif, a branch
CC   site acceptor, and a nucleophilic group for transesterifying a
CC   phosphodiester bond of an RNA. Together the two segments form an
CC   autocatalytic Y-branched intron which catalyses integration of at least
CC   the first segment into substrate RNA by a reverse splicing reaction
CC   The reverse-splicing introns are used, by specific cleavage and ligation
CC   of discontinuous nucleic acid, to generate new genes and gene products,
CC   e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC   manipulation, e.g. replacements for restriction enzymes) or
CC   immunologically active or signal-transducing proteins such as antibody
CC   and enzyme libraries.
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match            100.0%;  Score 1065;  DB 19;  Length 2162;
Best Local Similarity  100.0%;  Pred. No. 1.1e-236;
Matches 1065;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAAATGGGTCAGCTACCGTGGCACGCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 TCTGAGGGAAACAGTGACTGCTACTTTGGGAAATGGGTCAGCTACCGTGGCACGCACAGC 765

Qy     61 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 825

Qy    121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCCAAACATAATTACTGCCGG 885

Qy    181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG 945

Qy    241 GAGTACTGTGATGTGCCCTCGCTCCTGCTCCACCGTGGGCGGCCCTGAGACAGCAGTACAGCCAGCCTCAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 GAGTACTGTGATGTGCCCTCGCTCCTGCTCCACCGTGGGCGGCCCTGAGACAGTACAGCCAGCCTCAG 1005

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 1065

Qy    361 TTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 TTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1125

Qy    421 TCCTGCTGGATTCTCTCTGCCCACTGCTTCCGGGAGAGGTTTCCGCCCCACCACCTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 TCCTGCTGGATTCTCTCTGCCCACTGCTTCCGGGAGAGGTTTCCGCCCCACCACCTG 1185

Qy    481 ACGGTGATCTTGGCAGAACATACCGGGGTGGTCCCTGGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 ACGGTGATCTTGGCAGAACATACCGGGGTGGTCCCTGGGGTGGTCCCTGCGCAGGAGGAGCAGAAATTTGAA 1245

Qy    541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1305

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGCTGTCGCCAGGAGAGCAGCGTGGTTCCGCACT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 CTGCTGCGAGCTGAAATCGGATTCGTCCCGGCTGTCGCCAGGAGAGCAGCGTGGTTCCGCACT 1365

Qy    661 GTGTGCCTTCCCCGGCGACCTTGCCAGCTGCCGACTGACGGAGTGTGAGCTCTCCGGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 GTGTGCCTTCCCCCGGCGACCTTGCCAGCTGCCGACTGACGGAGTGTGAGCTCTCCGGC 1425

Qy    721 TACGGCAAGCAATGAGGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 TACGGCAAGCAATGAGGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1485

Qy    781 AGACTGCTACCCATCCAGCGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 AGACTGCTACCCATCCAGCGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC 1545

Qy    841 AACATGCTGTGTCGTGGAGACACTCGGAGCGCGGCCCCAGGCAAACTTGCACGACGCGC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 AACATGCTGTGTCGTGGAGACACTCGGAGCGCGGCCCCAGGCAAACTTGCACGACGCGCC 1605

Qy    901 TGCCAGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCCGACTTTGGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 TGCCAGGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACTTTGGTG 1665

Qy    961 GGCATCATCAGCTGGGACTCGGGGCCTGGCCTTCGTGGACAGAAGGATGTCCCGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 GGCATCATCAGCTGGGGGCCTGGGCCTTCGTGGACAGAAGGATGTCCCGGTGTGTACACAAAG 1725

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA 1065
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1770


RESULT 14
AAN60659
ID   AAN60659 standard; DNA; 2547 BP.
XX
AC   AAN60659;
XX
DT   25-MAR-2003 (updated)
DT   22-JUL-1991 (first entry)
XX
DE   Sequence encoding human-pre-tissue plasminogen activator (pre-t-PA).
XX
KW   Plasminogen conversion; vascular disease therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key          Location/Qualifiers
FT   CDS          87..191
FT                /*tag= a
FT   mat_peptide  192..1775
FT                /*tag= b
XX
PN   GB2173804-A.
XX
PD   22-OCT-1986.
XX
PF   21-APR-1986;  86GB-0609683.
XX
PR   22-APR-1985;  85US-0725468.
PR   01-APR-1986;  86US-0846697.
PR   22-APR-1985;  85US-0725468.
PR   01-APR-1986;  86US-0846697.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Heyneker HL, Vehar GA;
XX
DR   WPI; 1986-280715/43.
DR   P-PSDB; AAP60790.
XX
PT   New mutant forms of human tissue plasminogen activator- having
PT   higher specific activity and resistance to conversion to two chain
PT   form
XX
PS   Example; Fig 2; 34pp; English.
XX
CC   The patentors claim a novel recombinant human t-PA which is
CC   resistant to specific enzymatic cleavage because it is stabilised by
CC   site-directed mutagenesis at a 2-chain cleavage site. Partic. the
CC   natural Arg at position 275 is replaced by Gly or Glu, or Ile or at
```

```
CC   position 276 is replaced (275 and 276 refer to the posn. of the AAs
CC   in the mature protein; i.e. AAs 310 and 311 of AAP60790). Also new are
CC   (1) DNA sequences; (2) expression vectors; and (3) microorganisms
CC   and cell cultures transformed with these vectors.
CC   (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;

Query Match      100.0%;  Score 1065;  DB 7;  Length 2547;
Best Local Similarity   100.0%;  Pred. No.1.1e-236;
Matches 1065;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      711 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC 770

Qy       61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATCCATGATCCTGATAGGCAAGGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      771 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATCCATGATCCTGATAGGCAAGGTT 830

Qy      121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTGGCAAACATAATTACTGCCGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      831 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTGGCAAACATAATTACTGCCGG 890

Qy      181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      891 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 950

Qy      241 GAGTACTGTGATGTGCCCTCCTGCTCACCTGGGGCTTGAGACAGTACAGCCAGCCTCAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      951 GAGTACTGTGATGTGCCCTCCTGCTCACCTGGGGCTTGAGACAGTACAGCCAGCCTCAG 1010

Qy      301 TTTCGGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCACCCTGCAGCTGCCATC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1011 TTTCGGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCACCCTGCAGCTGCCATC 1070

Qy      361 TTTGGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1071 TTTGGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1130

Qy      421 TCCTGCTGGATTCTCTCTGCCAGACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1131 TCCTGCTGGATTCTCTCTGCCAGACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 1190

Qy      481 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1191 ACGGTGATCTTGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAATTTGAA 1250

Qy      541 GTCGAAAATACATTGTCCATAAGGAATTCGATAAGGAACTTACGACAACAATGACATTGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1251 GTCGAAAATACATTGTCCATAAGGAATTCGATAAGGAACTTACGACAACAATGACATTGCG 1310

Qy      601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGTGCCCGTGTCGTGCCAGGAGCAGCGGTCCGCACT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1311 CTGCTGCGAGCTGAAATCGGATTCGTCCCGGTGCCCGTGTCGTGCCAGGAGCAGCGGTCCGCACT 1370

Qy      661 GTGTGCCTTCCCCGGCGGACCTTCGACCTGGACTGGACTGGACGTGAGCGTGAGCGTGAGCTCTCCGGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1371 GTGTGCCTTCCCCGGCGGACCTTCGACCTGGACTGGACTGGACGTGAGCGTGAGCGTGAGCTCTCCGGC 1430

Qy      721 TACGGCAAGCATGAGGCCCTTGTCCTTTCATTCGGAGGCGGCTGAAGGAGGCTCATGTC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1431 TACGGCAAGCATGAGGCCCTTGTCCTTTCATTCGGAGGCGGCTGAAGGAGGCTCATGTC 1490

Qy      781 AGACTGTACCCATCCAGCCGCTGCACATCACAACACATTACTTAACAGAACAGTCACCGAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1491 AGACTGTACCCATCCAGCCGCTGCACATCACAACACATTACTTAACAGAACAGTCACCGAC 1550

Qy      841 AACATGCTGTGTCTGACACACTCGGAGGCGGCGGCCCCAAACTTGCGACGACGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1551 AACATGCTGTGTCTGACACACTCGGAGGCGGCGGCCCCAAACTTGCGACGACGCC 1610

Qy      901 TGCCAGGGGCGGATTCGGGGAGGGCCCCTGGTGTGTCTGAACGATGGCCGCATGAGTAGGCAAGGTT 960
```

```
Db     1611 TGCCAGGGCGGATTCGGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      961 GGCATCATCAGCTGGGCGTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTACACAAAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1671 GGCATCATCAGCTGGGCGTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTACACAAAG 1730

Qy     1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
              |||||||||||||||||||||||||||||||||||||||||||||
Db     1731 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1775

RESULT 15
AAN90542
ID   AAN90542 standard; cDNA; 2560 BP.
XX
AC   AAN90542;
XX
DT   25-MAR-2003  (updated)
DT   05-JUN-1990  (first entry)
XX
DE   Plasmid pKG12 contg. DNA encoding human melanoma t-PA.
XX
KW   Human tissue plasminogen-activator gene; pKG12; ss.
XX
OS   Homo sapiens.
XX
FH                Key             Location/Qualifiers
FT   sig_peptide  102..206
FT                /*tag= a
FT   mat_peptide  ..1787
FT                /*tag= b
FT                /product=t-PA
XX
PN   EP297066-A.
XX
PD   28-DEC-1988.
XX
PF   14-JUN-1988;  88EP-0850207.
XX
PR   18-JUN-1987;  87SE-0002562.
XX
PA   (KABI ) KABIGEN AB.
XX
PI   Pohl G, Hansson L, Loewenadler B;
XX
DR   WPI; 1989-001503/01.
DR   P-PSDB; AAP93716.
XX
PT   Modified tissue plasminogen-activator - having domains deleted and amino
PT   acid changes to increase biological half-life and reduce inactivation.
XX
PS   Disclosure; Fig 1; 18pp; English.
XX
CC   The sequence was obtained from cDNA prepd. from mRNA extracted from
CC   Bowes melanoma cells. The cDNA was used to construct a gene library which
CC   was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated
CC   and shown to comprise the whole coding region for human t-PA as well as
CC   102bp 5'flanking, 760 bp 3' flanking DNA and a poly A tail.
CC   See also AAN91608.
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;

Query Match      100.0%;  Score 1065;  DB 10;  Length 2560;
Best Local Similarity   100.0%;  Pred. No.1.1e-236;
Matches 1065;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      726 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACCGTGGCACGCACAGC 785

Qy       61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
```

`Wed Aug 20 09:29:40 2003

```
Db  786 CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGTT 845
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGG 180
        |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  846 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 905

Qy  181 AATCCTGATGGGATGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  906 AATCCTGATGGGATGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 965

Qy  241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGCCAGCCTCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  966 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGCCAGCCCTCAG 1025

Qy  301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1026 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 1085

Qy  361 TTTGCCAAGCACAGGAGGAGTGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1086 TTTGCCAAGCACAGGAGGAGTGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 1145

Qy  421 TCCTGCTGGATTCTCTCTGCGCCGCCACTGCTTCCAGGAGAGGTTCCGCCCACCACCTG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1146 TCCTGCTGGATTCTCTCTGCGCCGCCACTGCTTCCAGGAGAGGTTCCGCCCCACCACCTG 1205

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGTCCCTGCGGAGGAGCAGAAATTTGAA 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1206 ACGGTGATCTTGGGCAGAACATACCGGGTGTCCCTGCGGAGGAGCAGAAATTTGAA 1265

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGGACAATGACATTGCG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1266 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGGACAATGACATTGCG 1325

Qy  601 CTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCAGCGCTGGTCCGCACT 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1326 CTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCAGCGCTGGTCCGCACT 1385

Qy  661 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1386 GTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC 1445

Qy  721 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1446 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1505

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGAACAGTCACCGAC 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1506 AGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGAACAGTCACCGAC 1565

Qy  841 AACATGCTGTGTGCTGGAGACACTCGGAGACGGCGGGCCCCAGGCAAACTTGCACGACGCC 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1566 AACATGCTGTGTGCTGGAGACACTCGGAGACGGCGGGCCCCAGGCAAACTTGCACGACGCC 1625

Qy  901 TGCCAGGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1626 TGCCAGGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1685

Qy  961 GGCATCATCAGCTGGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1686 GGCATCATCAGCTGGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1745

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCCGTGA 1065
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1746 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCCGTGA 1790
```

Search completed: August 18, 2003, 23:28:51
Job time : 377.192 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     August 18, 2003, 22:51:23 ; Search time 83.3499 Seconds
                                (without alignments)
                            5639.757 Million cell updates/sec

Title:          US-09-987-455-4
Perfect score:  1065
Sequence:       1 tctgaggggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters:      1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Issued_Patents NA:*
               1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
               2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
               3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
               4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
               5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
               6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1065	100.0	1068	2	US-08-811-949-44	Sequence 44, Appl
2	1065	100.0	1419	2	US-08-811-949-62	Sequence 62, Appl
3	1065	100.0	1848	3	US-08-814-412-10	Sequence 10, Appl
4	1065	100.0	2101	2	US-08-811-949-42	Sequence 42, Appl
5	1065	100.0	2162	1	US-08-119-512-3	Sequence 3, Appli
6	1065	100.0	2162	1	US-08-488-015B-3	Sequence 3, Appli
7	1065	100.0	2162	1	US-08-488-015B-25	Sequence 25, Appl
8	1063.4	99.8	1170	2	US-08-811-949-64	Sequence 64, Appl
9	1063.4	99.8	1314	2	US-08-811-949-48	Sequence 48, Appl
10	1063.4	99.8	1738	6	5200340-1	Patent No. 5200340
11	1063.4	99.8	1955	2	US-08-883-795A-39	Sequence 39, Appl
12	1063.4	99.8	2457	6	5344773-1	Patent No. 5344773
13	1063.4	99.8	7360	5	US-08-286-740-1	Sequence 1, Appli
14	1063.4	99.8	7360	5	PCT-US95-09576-1	Sequence 1, Appli
15	1061.8	99.7	1068	2	US-08-811-949-46	Sequence 46, Appl
16	1061.8	99.7	1314	2	US-08-811-949-50	Sequence 50, Appl
17	1061.8	99.7	1974	2	US-08-811-949-38	Sequence 38, Appl
18	1057	99.2	1068	2	US-08-811-949-52	Sequence 52, Appl
19	1057	99.2	1314	2	US-08-811-949-54	Sequence 54, Appl
20	1056.2	99.2	1068	1	US-08-137-116-2	Sequence 2, Appli
21	1056.2	99.2	1137	4	US-09-553-498-9	Sequence 9, Appli
22	1056.2	99.2	1137	4	US-09-618-869-9	Sequence 9, Appli
23	1055.4	99.1	1170	2	US-08-811-949-66	Sequence 66, Appli
24	1054.8	99.0	1065	1	US-08-427-640-1	Sequence 1, Appli
25	1053.8	98.9	1314	2	US-08-811-949-56	Sequence 56, Appl
26	1050	98.6	1068	6	5223256-3	Patent No. 5223256
27	1046.8	98.3	1065	1	US-08-427-640-5	Sequence 5, Appli
28	1041	97.7	1068	2	US-08-811-949-58	Sequence 58, Appl
29	1035.4	97.2	1068	1	US-08-427-640-3	Sequence 3, Appli
30	1033	97.0	1065	2	US-08-811-949-60	Sequence 60, Appl
31	953.4	89.5	2497	6	5185259-2	Patent No. 5185259
32	924	86.8	1068	2	US-08-427-640-7	Sequence 7, Appli
33	801	75.2	1163	2	US-08-558-269-5	Sequence 5, Appli
34	801	75.2	1163	3	US-09-410-882-5	Sequence 5, Appli
35	626.6	58.8	1727	6	5244676-4	Patent No. 5244676
36	625	58.7	1605	6	5244676-1	Patent No. 5244676
37	472	44.3	472	2	US-08-811-949-40	Sequence 40, Appl
38	417	39.2	453	6	5200340-3	Patent No. 5200340
39	264.4	24.8	1724	6	5200340-5	Patent No. 5200340
40	170.8	16.0	1236	1	US-07-957-039A-7	Sequence 7, Appli
41	170.8	16.0	1475	4	US-09-643-597-122	Sequence 122, App
42	170.8	16.0	1475	4	US-09-480-884A-122	Sequence 122, App
43	170.8	16.0	1475	4	US-09-542-615A-122	Sequence 122, App
44	170.8	16.0	1475	4	US-09-606-421B-122	Sequence 122, App
45	170.8	16.0	2301	6	5188829-2	Patent No. 5188829

```
                              ALIGNMENTS

RESULT 1
US-08-811-949-44
; Sequence 44, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;  APPLICANT:  NIWA, MINEO
;  APPLICANT:  SAITO, YOSHIMASA
;  APPLICANT:  SASAKI, HITOSHI
;  APPLICANT:  HAYASHI, MASAKO
;  APPLICANT:  NOTANI, JOUJI
;  APPLICANT:  KOBAYASHI, MASAKAZU
;  TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;  NUMBER OF SEQUENCES: 67
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE:  P.C.
;   STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY:  ARLINGTON
;   STATE:  VA
;   COUNTRY:  USA
;   ZIP:  22202
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0,  Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/811,949
;   FILING DATE:  05-MAR-1997
;   CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;   NAME:  OBLON, NORMAN F.
;   REGISTRATION NUMBER:  24,618
;   REFERENCE/DOCKET NUMBER:  18-966-0
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  703-413-3000
;   TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  44:
;   SEQUENCE CHARACTERISTICS:
;    LENGTH:  1068 base pairs
;    TYPE:  nucleic acid
;    STRANDEDNESS:  double
;    TOPOLOGY:  circular
;   MOLECULE TYPE:  DNA (genomic)
;   FEATURE:
;    NAME/KEY:  CDS
;    LOCATION:  1..1065
US-08-811-949-44

Query Match             100.0%;  Score 1065;  DB 2;  Length 1068;
```

Best Local Similarity 100.0%; Pred. No. 1.4e-291; Indels 0; Gaps 0;
Matches 1065; Conservative 0; Mismatches 0;

```
Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACCGACAGC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACCGACAGC   63

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGACACTGGGCCTGGCAAACATAATTACTGCCGG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGACACTGGGCCTGGCAAACATAATTACTGCCGG  123

Qy    121  TACAACAGACAGAGACCCCAGTGCCAAGCCCTGGTGCGCAGGCTGCTGAAGAACCGCAGGCTGACGTGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124  TACAACAGACAGAGACCCCAGTGCCAAGCCCTGGTGCGCAGGCTGCTGAAGAACCGCAGGCTGACGTGG  183

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAGAGACAGTAGCAGCCAGCCTCAG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAGAGACAGTAGCAGCCAGCCTCAG  243

Qy    241  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTGACAGTAGCAGCCAGCCTCAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTGACAGTAGCAGCCAGCCTCAG  303

Qy    301  TTTCGCATCAAAGAGAGGGGCTCTTCGCCGACATCGCCTCGCCACCCCTGGCAGGCTGCCATC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304  TTTCGCATCAAAGAGAGGGGCTCTTCGCCGACATCGCCTCGCCACCCCTGGCAGGCTGCCATC  363

Qy    361  TTTGCCAAGCACAGAGAGGTCGCCCGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364  TTTGCCAAGCACAGAGAGGTCGCCCGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC  423

Qy    421  TCCTGCTGATTCTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCACCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424  TCCTGCTGATTCTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCACCTG  483

Qy    481  ACGGTGATCTTGGGCAGCAGAACATACCGGGTGGTCCCTGGCACGGGTGAGGAGGAGAAATTTGAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484  ACGGTGATCTTGGGCAGCAGAACATACCGGGTGGTCCCTGGCACGGGTGAGGAGGAGAAATTTGAA  543

Qy    541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  603

Qy    601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCGTGGTCCGACT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCGTGGTCCGACT  663

Qy    661  GTGTGCCCTTCCCCCGGCCCCCCGGCGGACCTGCAGCTGCCGGACTGCGCCGGAGTGTGAGCTCTCCCGGC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664  GTGTGCCCTTCCCCCGGCCCCCCGGCGGACCTGCAGCTGCCGGACTGCGCCGGAGTGTGAGCTCTCCCGGC  723

Qy    721  TACGGCAAGCATGAGGGCCCTTGTCTCCTTTCTATTCGGAAGGAGCGGCTGAAGGAGGCTCATGTC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724  TACGGCAAGCATGAGGGCCCTTGTCTCCTTTCTATTCGGAAGGAGCGGCTGAAGGAGGCTCATGTC  783

Qy    781  AGACTGTACCCATCCAGCCGCTGCACATCCACACATTTACTTAACAGAACAGTCACCGAC  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784  AGACTGTACCCATCCAGCCGCTGCACATCCACACATTTACTTAACAGAACAGTCACCGAC  843

Qy    841  AACATGCTGTGTCTGAGACACTCGGAGCGCGGGGCCCCAGGCAAACTTCACGACGCC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844  AACATGCTGTGTCTGAGACACTCGGAGCGCGGGGCCCCAGGCAAACTTCACGACGCC  903

Qy    901  TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGTG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904  TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGTG  963

Qy    961  GGCATCATCAGCTGGGGCTGGGGCTGGGCTGTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTACACAAAG  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964  GGCATCATCAGCTGGGGCTGGGGCTGGGCTGTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTACACAAAG  1023

Qy   1021  GTTACCAACTACCTAGACTGGATTCGTTGACAACATGCCGACCGTGA  1065
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1024  GTTACCAACTACCTAGACTGGATTCGTTGACAACATGCCGACCGTGA  1068
```

RESULT 2

US-08-811-949-62

; Sequence 62, Application US/08811949
; Patent No. 5840533
GENERAL INFORMATION:
    APPLICANT: NIWA, MINEO
    APPLICANT: SAITO, YOSHIMASA
    APPLICANT: SASAKI, HITOSHI
    APPLICANT: HAYASHI, MASAKO
    APPLICANT: NOTANI, JOUJI
    APPLICANT: KOBAYASHI, MASAKAZU
    TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
    NUMBER OF SEQUENCES: 67
    CORRESPONDENCE ADDRESS:
    ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, .
    ADDRESSEE: P.C.
    STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
    CITY: ARLINGTON
    STATE: VA
    COUNTRY: USA
    ZIP: 22202
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/811,949
    FILING DATE: 05-MAR-1997
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: OBLON, NORMAN F.
    REGISTRATION NUMBER: 24,618
    REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 703-413-3000
    TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
    LENGTH: 1419 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: circular
    MOLECULE TYPE: DNA (genomic)
FEATURE:
    NAME/KEY: CDS
    LOCATION: 1..1416
US-08-811-949-62

Query Match 100.0%; Score 1065; DB 2; Length 1419;
Best Local Similarity 100.0%; Pred. No. 1.5e-291;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACCGACAGC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    355  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACCGACAGC   414

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    415  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGATCCTGATAGGCAAGGTT  474

Qy    121  TACAACAGACAGAGACCCCAGTGCCAAGCCCTGGGCAAACATAATTACTGCCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    475  TACAACAGACAGAGACCCCAGTGCCAAGCCCTGGGCAAACATAATTACTGCCGG  534

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    535  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  594

Qy    241  GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGCCTGAGAGTAGCAGCCAGCCTCAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/814,412
;  FILING DATE:  11-MAR-1997
;  CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  Jarrell Ph.D., Brenda H.
;  REGISTRATION NUMBER:  39,223
;  REFERENCE/DOCKET NUMBER:  0079571-0040
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  617 248 5000
;  TELEFAX:  617 248 4000
INFORMATION FOR SEQ ID NO:  10:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  1848 base pairs
;  TYPE:  nucleic acid
;  STRANDEDNESS:  double
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  cDNA
;  IMMEDIATE SOURCE:
;  CLONE:  amplified t-PA clone
US-08-814-412-10

Query Match          100.0%;  Score 1065;  DB 3;  Length 1848;
Best Local Similarity 100.0%;  Pred. No. 1.7e-291;
Matches 1065;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCAGCGTGGCACGCACAGC    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    696  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCAGCGTGGCACGCACAGC   755

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    756  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   815

Qy    121  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCAAACATAATTACTGCCGG   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    816  TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCCAAACATAATTACTGCCGG   875

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    876  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   935

Qy    241  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCCTGAGACAGTACAGCCAGCCTCAG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    936  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCCTGAGACAGTACAGCCAGCCTCAG   995

Qy    301  TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    996  TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC   1055

Qy    361  TTTGCCAAGCACAGGAGGTCGCCCGGAGGCGGTTCCGCCCACCACCTG    420
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1056  TTTGCCAAGCACAGGAGGTCGCCCGGAGGCGGTTCCGCCCACCACCTG   1115

Qy    421  TCCTGCTGGATTCTCTGCGCCGCCACCTG    480
           |||||||||||||||||||||||||||||
Db   1116  TCCTGCTGGATTCTCTGCGCCGCCACCTG   1175

Qy    481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTGAA    540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1176  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTGAA   1235

Qy    541  GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG    600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1236  GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   1295

Qy    601  CTGCTGCAGCTGAAATCGGATTCGTCCGCGTTCGTCCCAGGAGAGCAGCGGCGTGGTCCGCACT    660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1296  CTGCTGCAGCTGAAATCGGATTCGTCCGCGTTCGTCCCAGGAGAGCAGCGGCGTGGTCCGCACT   1355
```

```
Db    595  GAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCCTGAGACAGTACAGCCAGCCTCAG   654
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301  TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655  TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC   714

Qy    361  TTTGCCAAGCACAGGAGGTCGCCCGGAGGCGGTTCCGCCGGGGGCATACTCATCAGC   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715  TTTGCCAAGCACAGGAGGTCGCCCGGAGGCGGTTCCGCCGGGGGCATACTCATCAGC   774

Qy    421  TCCTGCTGGATTCTCTGCGCCGCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775  TCCTGCTGGATTCTCTGCGCCGCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG   834

Qy    481  ACGGTGATCTTGGGCAGAACATACGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    835  ACGGTGATCTTGGGCAGAACATACGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA   894

Qy    541  GTCGAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG   600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895  GTCGAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG   954

Qy    601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGTCCCGGAGCAGCGGTGGTCCGCACT   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGTCCCGGAGCAGCGGTGGTCCGCACT   1014

Qy    661  GTGTGCGCCTTCCCCCGGCGGCGGCACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015  GTGTGCGCCTTCCCCCGGCGGCGGCACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC   1074

Qy    721  TACGGCAAGCATGAGGCCTTGTGTCCTTCTTCTCATTCGAGCGGCTGAACGGAGCTCATGTC   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075  TACGGCAAGCATGAGGCCTTGTGTCCTTCTTCTCATTCGAGCGGCTGAACGGAGCTCATGTC   1134

Qy    781  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC   1194

Qy    841  AACATGCTGTGTCTGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCCC   900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195  AACATGCTGTGTCTGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCCC   1254

Qy    901  TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGCATGGACTTTGGTG   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255  TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGCATGGACTTTGGTG   1314

Qy    961  GGCATCATCAGCTGGGGCCTGGGCTGGGCCTGGGTGTCCGGGTGTGTACACAAAG   1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315  GGCATCATCAGCTGGGGCCTGGGCTGGGCCTGGGTGTCCGGGTGTGTACACAAAG   1374

Qy   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA   1065
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1375  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA   1419
```

```
RESULT 3
US-08-814-412-10
; Sequence 10, Application US/08814412
; Patent No. 6151041
; GENERAL INFORMATION:
;  APPLICANT:  Jarrell Ph.D., Kevin A.
;  TITLE OF INVENTION:  Intron-Mediated Techniques and Reagents
;  NUMBER OF SEQUENCES:  46
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Choate, Hall & Stewart
;  STREET:  53 State Street
;  CITY:  Boston
;  STATE:  MA
;  COUNTRY:  USA
;  ZIP:  02109
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Floppy disk
```

```
;    FEATURE:
;    NAME/KEY:   CDS
;    LOCATION:   25..1710
US-08-811-949-42

Query Match              100.0%;   Score 1065;  DB 2;  Length 2101;
Best Local Similarity    100.0%;   Pred. No. 1.8e-291;
Matches 1065;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACGCTACCGTGCACGCACAGC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTCACGCTACCGTGCACGCACAGC  708

Qy     61  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709  CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  768

Qy    121  TACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769  TACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG  828

Qy    181  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGCGTGACGTGG  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829  AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGCGTGACGTGG  888

Qy    241  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTGACAGCCAGCCTCAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889  GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTGACAGCCAGCCTCAG  948

Qy    301  TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    949  TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAGGCTGCCATC  1008

Qy    361  TTTGCCAAGCACAGAGGTCGCCGCAGGAGTCGCGCGGTTCGTGTGCGGGGGCATACTCATCAGC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009  TTTGCCAAGCACAGAGGTCGCCGCAGGAGTCGCGCGGTTCGTGTGCGGGGGCATACTCATCAGC  1068

Qy    421  TCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCCGCCCCACCACCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1069  TCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCCGCCCCACCACCTG  1128

Qy    481  ACGGTGATCTTGGCAGAACATACCGGGTGGTGTCCCTGCCTGGTGCTGGAGGAGGCAGAAATTTGAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129  ACGGTGATCTTGGCAGAACATACCGGGTGGTGTCCCTGCCTGGTGCTGGAGGAGGCAGAAATTTGAA  1188

Qy    541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACAATGACATTGCG  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACAATGACATTGCG  1248

Qy    601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGAGCAGCGTGGTTCCGCACT  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1249  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGAGCAGCGTGGTTCCGCACT  1308

Qy    661  GTGTGCCTTCCCCGGGGACCTGCAGCTGCGACTGGATGCGGGGGTGAGCGTCTCCGGC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1309  GTGTGCCTTCCCCGGGGACCTGCAGCTGCGACTGGATGCGGGGGTGAGCGTCTCCGGC  1368

Qy    721  TACGGCAAGCATGAGGGCCTTGCTTCCTTTCATTCGGAGGCGGCTGAAGGAGGCTCATGTC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369  TACGGCAAGCATGAGGGCCTTGCTTCCTTTCATTCGGAGGCGGCTGAAGGAGGCTCATGTC  1428

Qy    781  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  1488

Qy    841  AACATGCTGTGTCTGAGACACTCGGACGCGGCCCCGTGTGTGTCTGAACGATGGCCGATGACTTTGGTG  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489  AACATGCTGTGTCTGAGACACTCGGACGCGGCCCCGTGTGTGTCTGAACGATGGCCGATGACTTTGGTG  1548

Qy    901  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549  TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG  1608

Qy    661  GTGTGCCTTCCCCGGGGACCTGCAGCTGCGACTGGATGCGGGGGTGAGCGTCTCCGGC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1356  GTGTGCCTTCCCCGGGGACCTGCAGCTGCGACTGGATGCGGGGGTGAGCGTCTCCGGC  1415

Qy    721  TACGGCAAGCATGAGGGCCTTGTCTCCTTTCATTCGGAGGCGGCTGAAGGAGGCTCATGTC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1416  TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCTCATGTC  1475

Qy    781  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1476  AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  1535

Qy    841  AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCCAGCAAACTTGCACGACGCC  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1536  AACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCCAGCAAACTTGCACGACGCC  1595

Qy    901  TGCCAGGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1596  TGCCAGGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTTGGTG  1655

Qy    961  GGCATCATCAGCTGGGGCCTGGGGCTGTGTGGACAGAAGAAGGATGTCCCGGTGTGTGTACACAAAG  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1656  GGCATCATCAGCTGGGGCCTGGGGCTGTGTGGACAGAAGAAGGATGTCCCGGTGTGTGTACACAAAG  1715

Qy   1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1065
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1716  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1760
```

```
RESULT 4
US-08-811-949-42
;  Sequence 42, Application US/08811949
Patent No.5840533
;  GENERAL INFORMATION:
;     APPLICANT:  NIWA, MINEO
;     APPLICANT:  SAITO, YOSHIMASA
;     APPLICANT:  SASAKI, HITOSHI
;     APPLICANT:  HAYASHI, MASAKO
;     APPLICANT:  NOTANI, JOUJI
;     APPLICANT:  KOBAYASHI, MASAKAZU
;     TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;     ADDRESSEE:  P.C.
;     STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;     CITY:  ARLINGTON
;     STATE:  VA
;     COUNTRY:  USA
;     ZIP:  22202
;     COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/811,949
;     FILING DATE:  05-MAR-1997
;     CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;     NAME:  OBLON, NORMAN F.
;     REGISTRATION NUMBER:  24,618
;     REFERENCE/DOCKET NUMBER:  18-966-0
;     TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  703-413-3000
;     TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  42:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH:  2101 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  circular
;     MOLECULE TYPE:  DNA (genomic)
```

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCTACCGTGCACGCACAGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  706 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCTACCGTGCACGCACAGC 765

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCGCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  766 CTCACCGAGTCGGGTGCCTCCTGCCTCCGCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 825

Qy  121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCAAACATAATTACTGCCGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  826 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCAAACATAATTACTGCCGG 885

Qy  181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGACGGCTGACGTG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  886 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGACGGCTGACGTG 945

Qy  241 GAGTACTGTGATGTGCCCTGCTCACCTGGGCGGCTGGAGACAGTACAGCCAGCCTCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  946 GAGTACTGTGATGTGCCCTGCTCACCTGGGCGGCTGGAGACAGTACAGCCAGCCTCAG 1005

Qy  301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCAGGCTGCCATC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1006 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCAGGCTGCCATC 1065

Qy  361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCGTGTGCGGGGCATACTCATCAGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1066 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCGTGTGCGGGGCATACTCATCAGC 1125

Qy  421 TCCTGCTGGATTCTCTCGCCCACTGCTTCCAGGAGAGGTTCCGCCCCACCACCTG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1126 TCCTGCTGGATTCTCTCGCCCACTGCTTCCAGGAGAGGTTCCGCCCCACCACCTG 1185

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCAGGAGGAGCAGAAATTTGAA 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1186 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCAGGAGGAGCAGAAATTTGAA 1245

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1305

Qy  601 CTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1306 CTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACT 1365

Qy  661 GTGTGCCTTCCCCGGCCGGACCTGCAGCTGCGACTGACCTGACGAGCGTGAGCTTCCCGGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1366 GTGTGCCTTCCCCGGCCGGACCTGCAGCTGCGACTGACCTGACGAGCGTGAGCTTCCCGGC 1425

Qy  721 TACGGCAAGCATGGCGATTCGGAGGACGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAACGTGTC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1426 TACGGCAAGCATGGCGATTCGGAGGACGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAACGTGTC 1485

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGAC 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1486 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGAC 1545

Qy  841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1546 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1605

Qy  901 TGCCAGGCGGATTCGGGAGGCCCCCGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1606 TGCCAGGCGGATTCGGGAGGCCCCCGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1665

Qy  961 GGCATCATCAGCTGGGCGCCTGGGCGTGTGGACAAGGATGTCCCGGGTGTGTACACAAAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1609 GGCATCATCAGCTGGGCGCCTGGGCGTGTGGACAAGGATGTCCCGGGTGTGTACACAAAG 1668

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCACCGTGA 1065
        |||||||||||||||||||||||||||||||||||||||||||||
Db 1669 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCACCGTGA 1713
```

RESULT 5
US-08-119-512-3
; Sequence 3, Application US/08119512
; Patent No. 5498531
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,512
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUI-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 82..334
; OTHER INFORMATION: /product= "Signal Sequence and
; OTHER INFORMATION: Finger-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 335..447
; OTHER INFORMATION: /product= "EGF-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448..714
; OTHER INFORMATION: /product= "Kringle-1 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 715..972
; OTHER INFORMATION: /product= "Kringle-2 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 973..2162
; OTHER INFORMATION: /product= "Catalytic domain"
US-08-119-512-3

Query Match   100.0%; Score 1065; DB 1; Length 2162;
Best Local Similarity 100.0%; Pred. No. 1.8e-291;

```
Db  766 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  825
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCTCTGGGCAACACATAATTACTGCCGG  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  826 TACAACAGCACAGAACCCCAGTGCCAGGCACTGGGCTCTGGGCAACACATAATTACTGCCGG  885

Qy  181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  886 AATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGG  945

Qy  241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCGGCTGAGACAGTACAGCCAGCCTCAG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  946 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCGGCTGAGACAGTACAGCCAGCCTCAG 1005

Qy  301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1006 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 1065

Qy  361 TTTGCCAAGCACACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1066 TTTGCCAAGCACACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 1125

Qy  421 TCCTGCTGGGATTCTCTCTGCGCGCCCACTGCTTCCGCCCCCACCACCTG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1126 TCCTGCTGGGATTCTCTCTGCGCGCCCACTGCTTCCGCCCCCACCACCTG 1185

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1186 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAA 1245

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1305

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGAGCGAGCAGGTCGGTCCGCACT  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1306 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGAGCGAGCAGGTCGGTCCGCACT 1365

Qy  661 GTGTGCCTTCCCCCGGCGACCTGCAGCTGCAGCTGGACCTGCCGGACTCGCCGGC  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1366 GTGTGCCTTCCCCCGGCGACCTGCAGCTGCAGCTGGACCTGCCGGACTCGCCGGC 1425

Qy  721 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGGAGGCGGCTGAAGGAGGCTCATGTC  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1426 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGGAGGCGGCTGAAGGAGGCTCATGTC 1485

Qy  781 AGACTGTACCATCAGCCGCTGCACATCACACATTACTTAACAGAACAGTCACCGAC  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1486 AGACTGTACCATCAGCCGCTGCACATCACACATTACTTAACAGAACAGTCACCGAC 1545

Qy  841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCGCCCCAGGCAAACTTGCAGACGCC  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1546 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCGCCCCAGGCAAACTTGCAGACGCC 1605

Qy  901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGACGATGCCGCATGAGACTTTGGTG  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1606 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGACGATGCCGCATGAGACTTTGGTG 1665

Qy  961 GGCATCATCAGCTGGGGCTGGGCCTGGGTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1666 GGCATCATCAGCTGGGGCTGGGCCTGGGTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1725

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGGCACCGACCGTGA 1065
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGGCACCGACCGTGA 1770
```

RESULT 6
US-08-488-015B-3
; Sequence 3, Application US/08488015B
; Patent No. 5780272
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,015B
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUV-008.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 82..334
; OTHER INFORMATION: /product= "Signal Sequence and
; Patent No. 5780272
; OTHER INFORMATION: Finger-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 335..447
; OTHER INFORMATION: /product= "EGF-like domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448..714
; OTHER INFORMATION: /product= "Kringle-1 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 715..972
; OTHER INFORMATION: /product= "Kringle-2 domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 973..2162
; OTHER INFORMATION: /product= "Catalytic domain"
US-08-488-015B-3

Query Match   100.0%; Score 1065; DB 1; Length 2162;
Best Local Similarity   100.0%; Pred. No. 1.8e-291;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCGCCAGC  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  706 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCGCCAGC 765
Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
```

RESULT 7
US-08-488-015B-25
; sequence 25, Application US/08488015B
; Patent No. 5780272
; GENERAL INFORMATION:

```
;   APPLICANT: Jarrell, Kevin A.
;   TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
;   TITLE OF INVENTION: AND REAGENTS
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley, Hoag & Eliot
;   STREET: One Post Office Square
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII (text)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/488,015B
;   FILING DATE: 07-JUN-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Vincent, Matthew P.
;   REGISTRATION NUMBER: 36,709
;   REFERENCE/DOCKET NUMBER: HUV-008.02
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617) 832-1000
;   TELEFAX: (617) 832-7000
;   INFORMATION FOR SEQ ID NO: 25:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2162 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-488-015B-25

Query Match           100.0%;   Score 1065;   DB 1;   Length 2162;
Best Local Similarity 100.0%;   Pred. No. 1.8e-291;
Matches 1065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGGCACGCACAGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     706 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGGCACGCACAGC 765

Qy      61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     766 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 825

Qy     121 TACACAGCACAGAACCCCAGTGCCCAGGACTGGGCCTGGCAAACATATTACTGCCGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     826 TACACAGCACAGAACCCCAGTGCCCAGGACTGGGCCTGGCAAACATAATTACTGCCGG 885

Qy     181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     886 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 945

Qy     241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCAGTACAGCCAGCCTCAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     946 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCAGTACAGCCAGCCTCAG 1005

Qy     301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTGGCAGGCTGCCATC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1006 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTGGCAGGCTGCCATC 1065

Qy     361 TTTGCCAAGCACAGGAGTGCCCCGAGAGCGGTTCCTGCGGGGGCATACTCATCAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1066 TTTGCCAAGCACAGGAGTGCCCCGAGAGCGGTTCCTGCGGGGGCATACTCATCAGC 1125

Qy     421 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1126 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 1185

Qy     481 ACGGTGATCTTGGGCAGAACATACCGGGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1186 ACGGTGATCTTGGGCAGAACATACCGGGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 1245

Qy     541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACATTCGACAATGACATTGCG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1246 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1305

Qy     601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGGCGGTGGTCCGCACT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1306 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGGCGGTGGTCCGCACT 1365

Qy     661 GTGTGCCTTCCCCGGCGACTGCCAGTCGACGAGTGTGAGCTCTCCGGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1366 GTGTGCCTTCCCCGGCGACTGCCAGTCGACGAGTGTGAGCTCTCCGGC 1425

Qy     721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1426 TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 1485

Qy     781 AGACTGTACCCATCCAGCCGCTGCACATCACACACATTTACTAACAGAACAGTCACCGAC 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1486 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTTAACAGAACAGTCACCGAC 1545

Qy     841 AACATGCTGTGTCTGGAGACACTCGGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1546 AACATGCTGTGTCTGGAGACACTCGGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 1605

Qy     901 TGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTGGTG 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1606 TGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1665

Qy     961 GGCATCATCAGCTGGGGCCTGGGCTTGGCTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1666 GGCATCATCAGCTGGGGCCTGGGCTTGGCTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1725

Qy    1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1726 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1770
```

```
RESULT 8
US-08-811-949-64
;   Sequence 64, Application US/08811949
;   Patent No. 5840533
;   GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERRNCE/DOCKET NUMBER: 18-966-0
```

```
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  703-413-3000
;     TELEFAX:  703-413-2220
INFORMATION FOR SEQ ID NO:  64:
SEQUENCE CHARACTERISTICS:
  LENGTH:  1170 base pairs
  TYPE:  nucleic acid
  STRANDEDNESS:  double
  TOPOLOGY:  circular
MOLECULE TYPE:  DNA (genomic)
FEATURE:
  NAME/KEY:  CDS
  LOCATION:  1..1167
US-08-811-949-64
```

```
Query Match            99.8%;  Score 1063.4;  DB 2;  Length 1170;
Best Local Similarity  99.9%;  Pred. No. 4e-291;
Matches 1064;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGC    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGGCACGCACAGC   165

Qy    61 CTCACCGAGTCGGTGCTGCCCTGCCTCCCGTGCCTCCCGTGCCTGATAGCTCAAGTT   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 CTCACCGAGTCGGTGCTGCCCTGCCTCCCGTGCCTCCCGTGCCTGATAGGCAAGTT    225

Qy   121 TACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG    180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226 TACACAGCAGAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG   285

Qy   181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   345

Qy   241 GAGTACTCTGATGTGCCCTCCTGCTCCACCTGCCGGCCTGAGACAGTACGACCGCCTCAG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 GAGTACTCTGATGTGCCCTCCTGCTCCACCTGCCGGCCTGAGACAGTACGACCGCCTCAG   405

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCCAGGCTGCCATC   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCCAGGCTGCCATC   465

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCATACTCATCAGC   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCATACTCATCAGC   525

Qy   421 TCCTGCTGGATTCTCTGCGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 TCCTGCTGGATTCTCTGCGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG   585

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA     540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA     645

Qy   541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGACATTGCG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   646 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGACATTGCG   705

Qy   601 CTGCTGCAGCTGGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGGTGGTCCGCACT   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   706 CTGCTGCAGCTGGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGGTGGTCCGCACT   765

Qy   661 GTGTGCCTTCCCCGGCGGACCTGCCAGCTGACCGAGTGTGAGCTCTCCGGC         720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   766 GTGTGCCTTCCCCGGCGGACCTGCCAGCTGACCGAGTGTGAGCTCTCCGGC         825

Qy   721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTGTCTCTATTCGGAGCGGCGGCTGAGGAGGCTCATGTC   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   826 TACGGCAAGCATGAGGGCCTTGTCTCCTTTGTCTCTATTCGGAGCGGCGGCTGAGGAGGCTCATGTC   885

Qy   781 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC    840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   886 AGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC    945

Qy   841 AACATGCTGTGTCTGGAGACACTG                                     900
           ||||||||||||||||||||||||
Db   946 AACATGCTGTGTCTGGAGACACTG                                    1005

Qy   901 TGCCAGGGGCGGATTCGGGAGGCCCCC                                  960
           ||||||||||||||||||||||||||
Db  1006 TGCCAGGGGCGGATTCGGGAGGCCCCC                                 1065

Qy   961 GGGCATCATCAGCTGGGGCCTGGGCTGTGGCGACAAGGATGTCCCGGGTGTGTACACAAAG  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1066 GGCATCATCAGCTGGGGCCTGGGCTGTGGCGACAAGGATGTCCCGGGTGTGTACACAAAG  1125

Qy  1021 GTTACCAACTACCTAGACTGGAATTCGTTGCGACAACATGCGACCGTGA            1065
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1126 GTTACCAACTACCTAGACTGGAATTCGTTGCACAACATGCGACCGTGA            1170
```

```
RESULT 9
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT:  NIWA, MINEO
;   APPLICANT:  SAITO, YOSHIMASA
;   APPLICANT:  SASAKI, HITOSHI
;   APPLICANT:  HAYASHI, MASAKO
;   APPLICANT:  NOTANI, JOUJI
;   APPLICANT:  KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES:  67
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;     ADDRESSEE:  P.C.
;     STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;     CITY:  ARLINGTON
;     STATE:  VA
;     COUNTRY:  USA
;     ZIP:  22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/811,949
;     FILING DATE:  05-MAR-1997
;     CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  OBLON, NORMAN F.
;     REGISTRATION NUMBER:  24,618
;     REFERENCE/DOCKET NUMBER:  18-966-0
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  703-413-3000
;     TELEFAX:  703-413-2220
INFORMATION FOR SEQ ID NO:  48:
SEQUENCE CHARACTERISTICS:
  LENGTH:  1314 base pairs
  TYPE:  nucleic acid
  STRANDEDNESS:  double
  TOPOLOGY:  circular
MOLECULE TYPE:  DNA (genomic)
FEATURE:
  NAME/KEY:  CDS
  LOCATION:  1..1311
US-08-811-949-48
```

```
Query Match            99.8%;  Score 1063.4;  DB 2;  Length 1314;
Best Local Similarity  99.9%;  Pred. No. 4.2e-291;
Matches 1064;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy     1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTGCGCACCACAGC   60
```

5200340-1
;Patent No. 5200340
;APPLICANT: FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
;TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;
;NUMBER OF SEQUENCES: 34
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/53,412
;FILING DATE: 22-MAY-1987
;SEQ ID NO:1:
;LENGTH: 1738
5200340-1

Query Match   99.8%; Score 1063.4; DB 6; Length 1738;
Best Local Similarity 99.9%; Pred. No. 4.8e-291;
Matches 1064; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 640 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 699

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 700 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 759

Qy 121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 760 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 819

Qy 181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 820 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 879

Qy 241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTCAGACAGTACAGCCAGCCTCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 880 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTCAGACAGTACAGCCAGCCTCAG 939

Qy 301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCTGGCAGGCTGCCATC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 940 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCTGGCAGGCTGCCATC 999

Qy 361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGTGCCGGCCGACTCGACGGAGTGTGAGCTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1000 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGTGCCGGCCGACTCGACGGAGTGTGAGCTC 1059

Qy 421 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCGCCCCACCACCTG 480
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1060 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCGCCCCACCACCTG 1119

Qy 481 ACGGTGATCTTGGGCAGAACATACCGGTGGTTCCGGGAGGAGCAGAGAAATTTGAA 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1120 ACGGTGATCTTGGGCAGAACATACCGGTGGTTCCGGGAGGAGCAGAGAAATTTGAA 1179

Qy 541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1180 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCC 1239

Qy 601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCATCGGCCTGTGTCGGC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1240 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCATCGGCCTGTGTCGGC 1299

Qy 661 GTGTGCCTTCCCCGGCGCCGAGTGTGACGGCAAACTTCACGACGCC 720
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1300 GTGTGCCTTCCCCGGCGCCGAGTGTGACGGCAAACTTCACGACGCC 1359

Qy 721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTGTGGAGCGGCTGAAGGAGGCTCATGTC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1360 TACGGCAAGCATGAGGGCCTTGTCTCCTTTGTGGAGCGGCTGAAGGAGGCTCATGTC 1419

Qy 781 AGACTGTACCCATCCAGCCGCCTGCACATTTACTTAACAGAACAGTCACCGAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1420 AGACTGTACCCATCCAGCCGCCTGCACATTTACTTAACAGAACAGTCACCGAC 1479

Db 250 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC 309

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 CTCACCGAGTCGGGTGCCTCCTGCCTCCTGCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 369

Qy 121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 370 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 429

Qy 181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 430 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 489

Qy 241 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTCAGACAGTACAGCCAGCCTCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 490 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTCAGACAGTACAGCCAGCCTCAG 549

Qy 301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCTGGCAGGCTGCCATC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 550 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCTGGCAGGCTGCCATC 609

Qy 361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGTGCCGGCCGACTCGACGGAGTGTGAGCTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 610 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGTGCCGGCCGACTCGACGGAGTGTGAGCTC 669

Qy 421 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCGCCCCACCACCTG 480
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 670 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCGCCCCACCACCTG 729

Qy 481 ACGGTGATCTTGGGCAGAACATACCGGTGGTTCCGGGAGGAGCAGAGAAATTTGAA 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 730 ACGGTGATCTTGGGCAGAACATACCGGTGGTTCCGGGAGGAGCAGAGAAATTTGAA 789

Qy 541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 790 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 849

Qy 601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCATCGGCCTGTGTCGGC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 850 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCATCGGCCTGTGTCGGC 909

Qy 661 GTGTGCCTTCCCCGGCGCCGAGTGTGACGGCAAACTTCACGACGCC 720
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 910 GTGTGCCTTCCCCGGCGCCGAGTGTGACGGCAAACTTCACGACGCC 969

Qy 721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTGTGGAGCGGCTGAAGGAGGCTCATGTG 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 970 TACGGCAAGCATGAGGGCCTTGTCTCCTTTGTGGAGCGGCTGAAGGAGGCTCATGTG 1029

Qy 781 AGACTGTACCCATCCAGCCGCCTGCACATTTACTTAACAGAACAGTCACCGAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1030 AGACTGTACCCATCCAGCCGCCTGCACATTTACTTAACAGAACAGTCACCGAC 1089

Qy 841 AACAGCTGTGTGCTGAGACACTCGGAGACGCGGGGGCCCCCAGGCAAACTTGACGAGCC 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1090 AACAGCTGTGTGCTGAGACACTCGGAGACGCGGGGGCCCCCAGGCAAACTTGACGAGCC 1149

Qy 901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGATGACTTTGGTG 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1150 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGATGACTTTGGTG 1209

Qy 961 GGCATCATCAGCTGGGGCCTGGGCCTGTGGACAGAAGGATGTCCCGGGTGTGTACAAAAG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1210 GGCATCATCAGCTGGGGCCTGGGCCTGTGGACAGAAGGATGTCCCGGGTGTGTACAAAAG 1269

Qy 1021 GTTACCAACTACCTTAGACTGGATTCGTGACAACATGCCAACCGTGA 1065
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 1270 GTTACCAACTACCTTAGACTGGATTCGTGACAACATGCCAACCGTGA 1314

RESULT 10

```
Qy   841  AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCGGGCCCCAGGCAAACTTGACGACGCC  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1480  AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCGGGCCCCAGGCAAACTTGACGACGCC 1539

Qy   901  TGCCAGGGCGGATTCGGGAGGCCCCTGGTGTGTCTGAACGATGCCGCATGACTTTGGTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1540  TGCCAGGGCGGATTCGGGAGGCCCCTGGTGTGTCTGAACGATGCCGCATGACTTTGGTG 1599

Qy   961  GGCATCATCAGCTCGGGTGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1600  GGCATCATCAGCTCGGGTGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1659

Qy  1021  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1660  GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1704
```

RESULT 11
US-08-883-795A-39
; Sequence 39, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
;   APPLICANT: Delcuve, Genevieve
;   APPLICANT: Awang, Gregor
;   TITLE OF INVENTION: Recombinant DNA Molecules and Expression
;   TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
;   NUMBER OF SEQUENCES: 39
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BERESKIN & PARR
;   STREET: 40 King Street West
;   CITY: Toronto
;   STATE: Ontario
;   COUNTRY: Canada
;   ZIP: M5H 3Y2
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/883,795A
;   FILING DATE: 27-JUN-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Gravelle, Micheline
;   REGISTRATION NUMBER: 40,261
;   REFERENCE/DOCKET NUMBER: 7841-062
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (416) 364-7311
;   TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 39:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1955 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
US-08-883-795A-39

Query Match   99.8%; Score 1063.4; DB 2; Length 1955;
Best Local Similarity   99.9%; Pred. No. 5e-291;
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy     1  TCTCAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  774

Qy    61  CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGTT  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   775  CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGTT  834
```

RESULT 12
5344773-1
;Patent No. 5344773
;   APPLICANT: WEI, CHA-MER;HSIUNG, NANCY;REDDY, VERMURI B.;
;LEMONTT, JEFFREY F.;DACKOWSKI, WILLIAM;DOUGLAS, RICHARD;
;COLE, EDWARD S.;PURCELL JR., RICHARD D.;LAU, DAVID TAI-YUI
;   TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN

(right column, continued alignment)

```
Qy   121  TACAACAGACAGAACCCCAGTGCCCAGGCACTCGGCCTTGGGCAAACATAATTACTGCCGG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835  TACAACAGACAGAACCCCAGTGCCCAGGCACTCGGCCTTGGGCAAACATAATTACTGCCGG  894

Qy   181  AATCCTGATGGGGATGGGCAAGCCCTGGTGCCAGTGTGCTGAAGAACCGAGGCTGACGTGG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895  AATCCTGATGGGGATGGGCAAGCCCTGGTGCCAGTGTGCTGAAGAACCGAGGCTGACGTGG  954

Qy   241  GAGTACGTGATGTGCCTCCACCTGCGCCTGAGACAGTACAGCCAGCCTCAG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955  GAGTACGTGATGTGCCTCCACCTGCGCCTGAGACAGTACAGCCAGCCTCAG 1014

Qy   301  TTTCGCATCAAAGGAGGGCTCTTGCCCGACATCGCCTCCCACCCCTGGCCAGGCTGCCATC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1015  TTTCGCATCAAAGGAGGGCTCTTGCCCGACATCGCCTCCCACCCCTGGCCAGGCTGCCATC 1074

Qy   361  TTTGCCAAGCACAGGAGGTCGCCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1075  TTTGCCAAGCACAGGAGGTCGCCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 1134

Qy   421  TCCTGCTGGATTCTTCTCTGCGCCCACTGCTTCCAGGAGGAGGTTCCGCCCCCACCACCTG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1135  TCCTGCTGGATTCTTCTCTGCGCCCACTGCTTCCAGGAGGAGGTTCCGCCCCCACCACCTG 1194

Qy   481  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1195  ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 1254

Qy   541  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTACGACAATGACAATTGCCG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255  GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTACGACAATGACAATTGCCG 1314

Qy   601  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGCGTGGTCCGCACT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315  CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGCGTGGTCCGCACT 1374

Qy   661  GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCCGGACTGGAGTGTGAGCTCTCCGGC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1375  GTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCCGGACTGGAGTGTGAGCTCTCCGGC 1434

Qy   721  TACGGCAAGCAATGCGAGGCCTTGTCTGTCCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435  TACGGCAAGCAATGCGAGGCCTTGTCTGTCCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1494

Qy   781  AGACTGTACCCATCCAGCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1495  AGACTGTACCCATCCAGCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 1554

Qy   841  AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCGGGCCCCAGGCAAACTTGCACGACGCC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1555  AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCGGGCCCCAGGCAAACTTGCACGACGCC 1614

Qy   901  TGCCAGGGCGGATTCGGGAGGCCCCTGGTGTCTGAACGATGATGCCGCATGACTTTGGTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1615  TGCCAGGGCGGATTCGGGAGGCCCCTGGTGTCTGAACGATGATGCCGCATGACTTTGGTG 1674

Qy   961  GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1675  GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTACACCAAG 1734

Qy  1021  GTTACCAACTACTTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1735  GTTACCAACTACCTAGACTGGATTCGTGACAACATGGACCGTGA 1779
```

;ACTIVATOR PRODUCED BY RECOMBIANT DNA
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/782,686
; FILING DATE: 01-OCT-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656,770
; FILING DATE: 01-OCT-1984
;SEQ ID NO:1:
; LENGTH: 2457
5344773-1

Query Match   99.8%;   Score 1063.4;   DB 6;   Length 2457;
Best Local Similarity   99.9%;   Pred. No. 5.5e-291;   Indels   0;   Gaps   0;
Matches 1064; Conservative   0; Mismatches   0;   Indels   1;   Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   638 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 697

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCGTCCGTGAATTCCATGATCCTGA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   698 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCGTCCGTGAATTCCATGATCCTGA 757

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   758 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 817

Qy   181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   818 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 877

Qy   241 GAGTACTGTGATGCCCCTCCTGCTCCACCTGGGCCTCGGGCTACAGCAGTACAGCCCAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   878 GAGTACTGTGATGCCCCTCCTGCTCCACCTGGGCCTCGGGCTACAGCAGTACAGCCTCAG 937

Qy   301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   938 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 997

Qy   361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   998 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC 1057

Qy   421 TCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCCGGCCCCCACCACCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1058 TCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCCGGCCCCCACCACCTG 1117

Qy   481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1118 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA 1177

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1178 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1237

Qy   601 CTGCTGCAGCTGAAATGGAATTCGTCCCGGCTTGTCTCCCGGCTGCCGCTGCCTGCCCGCT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1238 CTGCTGCAGCTGAAATGGAATTCGTCCCGGCTTGTCTCCCGGCTGCCGCTGCCTGCCCGCT 1297

Qy   661 GTGTGCCTTCCCCGGCGGGGCAGACTGCAGCTGCGGGACTGACGAGTGTGACGGACTCTCCGGC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1298 GTGTGCCTTCCCCGGCGGGGCAGACTGCAGCTGCGGGACTGACGAGTGTGACGGACTCTCCGGC 1357

Qy   721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1358 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1417

Qy   781 AGACTGTACCCATCCAGCCGCGTGCACATTCCACATTCACAACAGTCACCGAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1418 AGACTGTACCCATCCAGCCGCGTGCACATTCCACATTCACAACAGTCACCGAC 1477

Qy   841 AACATGCTGTGTCTGCGAGACACTCGGAGCACCGGAGGCGGCGGCGGGCGGG 900

Db   1478 AACATGCTGTGTCTGCGAGACACTCGGAGCACCGGAGGCGGCGGGCGGGCCCCAGGCAAACTTGCACGACGCC 1537
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   901 TGCCAGGCGATTCGGGAGGCGGAGGCCCCTGTGTGTCTGAACGATGCCGATGACTTGGTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1538 TGCCAGGCGATTCGGGAGGCGGAGGCCCCTGTGTGTCTGAACGATGCCGATGACTTGGTG 1597

Qy   961 GGCATCATCAGCTGGCCTGGGCCTGGGCTGTGCAGACAAGGATGTCCCGGGTGTGTACACAAAG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1598 GGCATCATCAGCTGGCCTGGGCCTGGGCTGTGCAGACAAGGATGTCCCGGGTGTGTACACCAAG 1657

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA 1065
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA 1702

RESULT 13
US-08-286-740-1
; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
; APPLICANT: Crowley, Craig W.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,740
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7360 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-286-740-1

Query Match   99.8%;   Score 1063.4;   DB 1;   Length 7360;
Best Local Similarity   99.9%;   Pred. No. 8.7e-291;   Indels   0;   Gaps   0;
Matches 1064; Conservative   0; Mismatches   0;   Indels   1;   Gaps   0;

Qy   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCCACAGC 2305

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCCTCCGTGAATTCCATGATCCTGATAGGCAAGGTT 2365

Qy   121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09576
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286740
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 798PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7360 bases
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US95-09576-1

Query Match   99.8%; Score 1063.4; DB 5; Length 7360;
Best Local Similarity 99.9%; Pred. No. 8.7e-291;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGTCAGCCTACCGTGCACCGCACAGC   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGTCAGCCTACCGTGCACGCACAGC  2305

Qy     61 CTCACCGAGTCGGGTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT         120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 CTCACCGAGTCGGGTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT        2365

Qy    121 TACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2366 TACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  2425

Qy    181 AATCCTGATGGGGATGGGATGCCAAGCCCTGTGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 AATCCTGATGGGGATGGGATGCCAAGCCCTGTGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  2485

Qy    241 GAGTACTGTGATGTGCCCGCTGCCTGCTGCCTGAGACAGTACAGCCAGCCTCAG               300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 GAGTACTGTGATGTGCCCGCTGCCTGCTGCCTGAGACAGTACAGCCAGCCTCAG              2545

Qy    301 TTTCGCATCAAAGGAGGGGCTCTTCGCCACGTGCTGCCATC                           360
          |||||||||||||||||||||||||||||||||||||||||
Db   2546 TTTCGCATCAAAGGAGGGGCTCTTCGCCACGTGCTGCCATC                          2605

Qy    361 TTTGCCAAGCACAGAGGAGGTCGCCCGGAGAGCGGTTCCTGTGTGCGGGGGCATACTCATCAGC    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 TTTGCCAAGCACAGAGGAGGTCGCCCGGAGAGCGGTTCCTGTGTGCGGGGGCATACTCATCAGC   2665

Qy    421 TCCTGCTGGATTCTCTGCCGCCGCACCTG                                       480
          |||||||||||||||||||||||||||||
Db   2666 TCCTGCTGGATTCTCTGCCGCCGCACCTG                                      2725

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGAGAGAAATTTGAA             540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2726 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGAGAGAAATTTGAA            2785
```

```
Db   2366 TACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  2425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 AATCCTGATGGGGATGGGATGCCAAGCCCTGTGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 AATCCTGATGGGGATGGGATGCCAAGCCCTGTGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG  2485
Qy    241 GAGTACTGTGATGTGCCCGCTGCCTGCTGCCTGAGACAGTACAGCCAGCCTCAG               300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 GAGTACTGTGATGTGCCCGCTGCCTGCTGCCTGAGACAGTACAGCCAGCCTCAG              2545
Qy    301 TTTCGCATCAAAGGAGGGGCTCTTCGCCACGTGCTGCCATC                           360
          |||||||||||||||||||||||||||||||||||||||||
Db   2546 TTTCGCATCAAAGGAGGGGCTCTTCGCCACGTGCTGCCATC                          2605
Qy    361 TTTGCCAAGCACAGAGGAGGTCGCCCGGAGAGCGGTTCCTGTGTGCGGGGGCATACTCATCAGC    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 TTTGCCAAGCACAGAGGAGGTCGCCCGGAGAGCGGTTCCTGTGTGCGGGGGCATACTCATCAGC   2665
Qy    421 TCCTGCTGGATTCTCTGCCGCCGCACCTG                                       480
          |||||||||||||||||||||||||||||
Db   2666 TCCTGCTGGATTCTCTGCCGCCGCACCTG                                      2725
Qy    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGAGAGAAATTTGAA             540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2726 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGAGAGAAATTTGAA            2785
Qy    541 GTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG         600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2786 GTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG        2845
Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGGCTGTGCCCGCTGTGCCCGCTG                 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 CTGCTGCAGCTGAAATCGGATTCGTCCCGGCTGTGCCCGCTGTGCCCGCTG                2905
Qy    661 GTGTGCCTTCCCCCGGCTGCCAGCTGCAGCTGGACCTGACCTGGACGAGTGTGAGCTCTCCGGC    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 GTGTGCCTTCCCCCGGCTGCCAGCTGCAGCTGGACCTGACCTGGACGAGTGTGAGCTCTCCGGC   2965
Qy    721 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCATTCGGAGCGCGGCTGGAGCGGAGGCTCATGTC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCATTCGGAGCGCGGCTGGAGCGGAGGCTCATGTC  3025
Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC        840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC       3085
Qy    841 AACATGTGTCGTGTGGAGACACTCGGAGCGCGGCCCCAGGCACAAACTTGCACGACGCC         900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 AACATGTGTCGTGTGGAGACACTCGGAGCGCGGCCCCAGGCACAAACTTGCACGACGCC        3145
Qy    901 TGCCAGGGGCGATTCGGGAGGCCCCCTGTGTGTCTGGAGCGATGGCCGCGATGGCCATGAGA     960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 TGCCAGGGGCGATTCGGGAGGCCCCCTGTGTGTCTGGAGCGATGGCCGCGATGGCCATGAGA    3205
Qy    961 GGCATCATCAGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG      1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3206 GGCATCATCAGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGTGTACACCAAG     3265
Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA                   1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTGA                  3310
```

RESULT 14
PCT-US95-09576-1
; Sequence 1, Application PC/TUS9509576
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```
Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2786 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   2845

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCCGCTGTCGCCGCTGGTCCGCACT            660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 CTGCTGCAGCTGAAATCGGATTCGTCCCCGCTGTCGCCGCTGGTCCGCACT           2905

Qy    661 GTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 GTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGC   2965

Qy    721 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   3025

Qy    781 AGACTGTACCCATCCAGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC    840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 AGACTGTACCCATCCAGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC   3085

Qy    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGAGCGCC    900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGAGCGCC   3145

Qy    901 TGGCAGGGGCGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 TGGCAGGGGCGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACTTTGGTG   3205

Qy    961 GGCATCATCAGCTGGGGCCTGGGCTGTGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   3206 GGCATCATCAGCTGGGGCCTGGGCTGTGTGACAGAAGGATGTCCCGGGTGTGTACACCAAG   3265

Qy   1021 GTTACCAACTAACCTAGACTGGATTCGTGACAACATGCGACCGTGA   1065
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 GTTACCAACTAACCTAGACTGGATTCGTGACAACATGCGACCGTGA   3310
```

RESULT 15
US-08-811-949-46
; Sequence 46, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1065
US-08-811-949-46

```
Query Match        99.7%;   Score 1061.8;  DB 2;  Length 1068;
Best Local Similarity  99.8%;   Pred. No. 1.1e-290;
Matches 1063;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TCTGAGGGAAACAGTGACTGCTACTTGGGAATGGGTCAGCCTACCGTGCAGCCACAGC    60
        |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCAGCCACAGC    63

Qy   61 CTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCAGGCACTGGCAATTCCATGATCCGATAGAGGCAAGGTT   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 CTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCAGGCACTGGCAATTCCATGATCCGATAGAGGCAAGGTT   123

Qy  121 TACACAGCACAGAACCCCAGTGCCAGTGGCCTGGCCTGGGCAAACATAATTACTGCCGG   180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124 TACACAGCACAGAACCCCAGTGCCAGTGGCCTGGCCTGGGCAAACATAATTACTGCCGG   183

Qy  181 AATCCTGATGGGATGCCAAGCCCTGGTGCCAGTGCTGAAGAACCCGGCTGACGTGG   240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184 AATCCTGATGGGATGCCAAGCCCTGGTGCCAGTGCTGAAGAACCCGGCTGACGTGG   243

Qy  241 GAGTACTGTGGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244 GAGTACTGTGGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG   303

Qy  301 TTTCGCATCAAAGGAGGGGCTCTTGCCGCACATGCCCGCAGCATGCTGCTGCCTCCACCCCTGGCAGGCGTGCCATC   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  304 TTTCGCATCATAGGAGGGGCTCTTGCCGCACATGCCCGCAGCATGCTGCTGCCTCCACCCCTGGCAGGCGTGCCATC   363

Qy  361 TTTGCCAAGCACAGGAGAGTCGCCCGGACTGCGGCGGTTCCTGTGTGCGGGGGCATACTCATCAGC   420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364 TTTGCCAAGCACAGGAGAGTCGCCCGGACTGCGGCGGTTCCTGTGTGCGGGGGCATACTCATCAGC   423

Qy  421 TCCTGCTGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG   480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424 TCCTGCTGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGTTTCCGCCCCACCACCTG   483

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA   540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA   543

Qy  541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG   603

Qy  601 CTGCTGCAGCTGAAATCGGATTCGTCCCCGGCGACCTGCAGCTGCCGGACTGGACGCGCT   660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604 CTGCTGCAGCTGAAATCGGATTCGTCCCCGGCGACCTGCAGCTGCCGGACTGGACGCGCT   663

Qy  661 GTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGCGCTGTGAGCTGTCCGGC   720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664 GTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGCGCTGTGAGCTGTCCGGC   723

Qy  721 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  724 TACGGCAAGCATGAGGCCTTGTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC   783

Qy  781 AGACTGTACCCATCCAGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC   840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784 AGACTGTACCCATCCAGCCGCTGCATCACACATTTACTTAACAGAACAGTCACCGAC   843
```

```
Qy   841 AACATGCTGTGTGTGCTGGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACTTGCACGAGCGCC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844 AACATGCTGTGTGTGCTGGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACTTGCACGAGCGCC  903

Qy   901 TGCCAGGGCGATTCGGGAGGCCCCCGTGTCTGTCTGAACGATGGCCGCATGACTTTGGTG  960
         |||||||||||||||||||||||||||||||--||||||||||||||||||||||||||
Db   904 TGCCAGGGCGATTCGGGAGGCCCCCGTGTGTCTGTCTGAACGATGGCCGCATGACTTTGGTG  963

Qy   961 GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964 GGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1023

Qy  1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1065
         |||||||||||||||||||||||||||||||||||||||||||
Db  1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA  1068
```

Search completed: August 19, 2003, 08:35:09
Job time : 84.3499 secs

OM nucleic - nucleic search, using sw model

```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: August 19, 2003, 04:46:44 ; Search time 300.529 Seconds
                                            (without alignments)
                                            7930.701 Million cell updates/sec

Title: US-09-987-455-4
Perfect score: 1065
Sequence: 1 tctgagggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY NUC
                    Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA:*
  1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
  2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
  3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
  4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
  5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
  6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
  7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
  8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
  9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1065	100.0	1065	11	US-09-987-455-4	Sequence 4, Appli
2	1065	100.0	1065	9	US-09-987-455-7	Sequence 7, Appli
3	1063.4	99.8	1689	9	US-09-969-271-6	Sequence 6, Appli
4	1063.4	99.8	2509	14	US-10-193-656-7	Sequence 7, Appli
5	1063.4	99.8	2519	9	US-09-969-271-5	Sequence 5, Appli
6	1062	99.7	1128	11	US-09-987-455-2	Sequence 2, Appli
7	1062	99.7	1128	11	US-09-987-455-5	Sequence 5, Appli
8	1061.8	99.7	2641	10	US-09-974-298-144	Sequence 144, App
9	230.2	21.6	329	12	US-10-007-926A-433	Sequence 433, App
10	191.6	18.0	1212	10	US-09-880-503-15	Sequence 15, Appl
11	185	17.4	1415	14	US-10-198-846-12748	Sequence 12748, A
12	170.8	16.0	1475	10	US-09-735-705-122	Sequence 122, App
13	170.8	16.0	1475	10	US-09-850-716A-122	Sequence 122, App
14	170.8	16.0	1475	10	US-09-897-778-122	Sequence 122, App
15	170.8	16.0	1475	11	US-09-466-396A-122	Sequence 122, App
16	170.8	16.0	1475	12	US-10-117-982-122	Sequence 122, App
17	170.8	16.0	1475	12	US-10-101-510-159	Sequence 159, App
18	170.8	16.0	2336	12	US-09-971-392-46	Sequence 46, Appl
19	170.8	16.0	2336	12	US-10-101-510-458	Sequence 458, App
20	167.6	15.7	1296	13	US-10-076-421-1	Sequence 1, Appli
21	167.6	15.7	2294	10	US-09-735-705-123	Sequence 123, App
22	167.6	15.7	2294	10	US-09-850-716A-123	Sequence 123, App
23	167.6	15.7	2294	11	US-09-897-778-123	Sequence 123, App
24	167.6	15.7	2294	12	US-09-466-396A-123	Sequence 123, App
25	167.6	15.7	2294	12	US-10-117-982-23	Sequence 123, App
26	167.6	15.7	2344	12	US-10-101-510-383	Sequence 383, App
27	166	15.6	474	9	US-09-864-761-1682	Sequence 1682, Ap
28	166	15.6	1236	10	US-09-880-503-12	Sequence 12, Appl
29	166	15.6	2304	12	US-10-301-822-160	Sequence 160, App
30	162	15.2	2304	14	US-10-171-311-183	Sequence 183, App
31	162	15.2	2486	14	US-10-106-698-1989	Sequence 1989, Ap
32	159.8	15.0	831	10	US-09-880-503-14	Sequence 14, Appl
33	159.8	15.0	972	10	US-09-880-503-16	Sequence 16, Appl
34	139	13.1	461	11	US-09-918-995-10405	Sequence 10405, A
35	123	11.5	1959	12	US-10-101-510-743	Sequence 743, App
36	123	11.5	2035	12	US-10-101-510-448	Sequence 448, App
37	123	11.5	2048	11	US-09-858-909-1	Sequence 1, Appli
38	120.4	11.3	2036	10	US-09-954-456-552	Sequence 552, App
39	120.4	11.3	2036	10	US-09-880-107-1612	Sequence 1612, Ap
c 40	114.4	10.7	116	9	US-09-864-761-18438	Sequence 18438, A
41	105	9.9	457	11	US-09-918-995-11188	Sequence 11188, A
c 42	92.4	8.7	372	10	US-09-920-300A-237	Sequence 237, App
c 43	92.4	8.7	372	13	US-10-033-528-237	Sequence 237, App
44	88	8.3	442	11	US-09-918-995-12715	Sequence 12715, A
45	83.8	7.9	1230	9	US-09-879-792-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-987-455-4
; Sequence 4, Application US/09987455
; Publication No. US2003049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00277779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-4

Query Match   100.0%; Score 1065; DB 11; Length 1065;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TCTGAGGGAAACAGTGACTTGGGAATGGTCAGCGTCACCGTGGCACCCACAGC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTGAGGGAAACAGTGACTTGGGAATGGTCAGCGTCACCGTGGCACCCACAGC 60

Qy     61 CTCACCGAGTCGGGGTGCCTCCCGTCCCGTGGAATTCCATGATCCTGGATAGGCAAGGTT 120
```

```
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
;   SEQ ID NO 7
;   LENGTH: 1065
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: coding
;   OTHER INFORMATION: sequence for K2S protein
US-09-987-455-7

Query Match            100.0%;   Score 1065;   DB 11;   Length 1065;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 TCTGAGGGAAACAGTGACTGACTTGGAATGGGTCAGTCGCACGCACCACAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTGAGGGAAACAGTGACTGACTTGGAATGGGTCAGTCGCACGCACCACAGC 60

Qy     61 CTCACCGAGTCGGGTGCCTCCCGCCTCCCGCCTCCGTGATCCTGATAGGCAAGGTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCACCGAGTCGGGTGCCTCCCGCCTCCCGCCTCCGTGATCCTGATAGGCAAGGTT 120

Qy    121 TACACAGCAGAACCCGAGGCACTGGGCTGGCAAACATAATTACTGCCGG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACAGCAGAACCCGAGGCACTGGGCTGGCAAACATAATTACTGCCGG 180

Qy    181 AATCCTGATGGGGATGCCCCTCCGCCTGAGAGAACCGCAGGCTGACGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATCCTGATGGGGATGCCCCTCCGCCTGAGAGAACCGCAGGCTGACGTGG 240

Qy    241 GAGTACTGTGATGCCTCCACCTGGGCCATGCAGGCTGCCATC 300
          ||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGTACTGTGATGCCTCCACCTGGGCCATGCAGGCTGCCATC 300

Qy    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCACCCCGCAGGCTGCCATC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCACCCCGCAGGCTGCCATC 360

Qy    361 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 420

Qy    421 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGGGGCATACTCATCAGC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGGGGCATACTCATCAGC 480

Qy    481 GTCGAAAATACATTGTCCATAAGGAATTCGACAATGACACTTACGACAATTGGCCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTCGAAAATACATTGTCCATAAGGAATTCGACAATGACACTTACGACAATTGGCCG 540

Qy    541 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCGCTGTGCCGAGAGCAGCGTCCGCACT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCGCTGTGCCGAGAGCAGCGTCCGCACT 600
```

RESULT 2
US-09-987-455-7
; Sequence 7, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:

```
Qy  301 TTTCGCATCAAAGGAGGGCTCTTCTGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 360
        ||||||||||||||||  |||||||| ||||||||||||||||||||||||||||||||||
Db  925 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC 984

Qy  361 TTTGCCAAGCACAGGAGGTCGCCCGGAGCGGTTCCTGTGCGGGGCCATACTCATCAGC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  985 TTTGCCAAGCACAGGAGGTCGCCCGGAGCGGTTCCTGTGCGGGGCCATACTCATCAGC 1044

Qy  421 TCTGCTGGATTCTCTCTGCCGCACTGCTTCCAGGAGAGTTCCGCCCCACCACCTG 480
        ||  |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 1045 TCCTGCTGGATTCTCTCTGCCGCACTGCTTCCAGGAGAGTTCCGCCCCACCACCTG 1104

Qy  481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCTGGTCCTGGCCGAGAAATTTGAA 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1105 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCTGGTCCTGGCCGAGAAATTTGAA 1164

Qy  541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1165 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 1224

Qy  601 CTTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGCTGGTCCGCACT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db 1225 CTTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGCGGTCCGCACT 1284

Qy  661 GTGTGCCTTCCCCCGGCCGACCTGCCAGCTGCCCGGACTGGAGTGTGAGCTCTCCGGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1285 GTGTGCCTTCCCCCGGCCGACCTGCCAGCTGCCCGGACTGGAGTGTGAGCTCTCCGGC 1344

Qy  721 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1345 TACGGCAAGCATGAGGCCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 1404

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTAACAGAACAGTCACCGAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1405 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTAACAGAACAGTCACCGAC 1464

Qy  841 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1465 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC 1524

Qy  901 TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1525 TGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1584

Qy  961 GGCATCATCAGCTGGGGCCTGGGCTGTGGAGACAGAAGGATGTCCCGGGTGTGTACACCAAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1585 GGCATCATCAGCTGGGGCCTGGGCTGTGGAGACAGAAGGATGTCCCGGGTGTGTACACCAAG 1644

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1645 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1689
```

```
Qy  661 GTGTGCCTTCCCCCGGCCGACCTGCAGCTGCGCGGACTGGAGTGTGAGCTCTCCGGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GTGTGCCTTCCCCCGGCCGACCTGCAGCTGCGCGGACTGGAGTGTGAGCTCTCCGGC 720

Qy  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 780

Qy  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTAACAGAACAGTCACCGAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTAACAGAACAGTCACCGAC 840

Qy  841 AACATGTCGTGTGAGACACTCGGAGCGCGGGGCCCCAGGCAAACTTGCACGACGCC 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 AACATGTCGTGTGAGACACTCGGAGCGCGGGGCCCCAGGCAAACTTGCACGACGCC 900

Qy  901 TGCCAGGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TGCCAGGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960

Qy  961 GGCATCATCAGCTGGGGCGTGGGGCCTGGGGCTGTGGACAGAAGAATGTCCCGGGTGTGTACACAAAG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 GGCATCATCAGCTGGGGCGTGGGGCCTGGGGCTGTGGACAGAAGAATGTCCCGGGTGTGTACACAAAG 1020

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCGTGA 1065
```

Query Match   99.8%; Score 1063.4; DB 9; Length 1689;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy    1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
        ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  625 TCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 684

Qy   61 CTCACCGAGTCGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  685 CTCACCGAGTCGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 744

Qy  121 TACACAGCACAGACCCCAGTGCCAGGCACTGGGCCTGGCCAGCAAACATAATTACTGCCGG 180
        ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  745 TACACAGCACAGACACCCCAGTGCCAGGCACTGGGCCTGGCCAGCAAACATAATTACTGCCGG 804

Qy  181 AATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG 240
        ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  805 AATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCCGCAGGCTGACGTGG 864

Qy  241 GAGTACTGTGATGTGCCCTCTGCTCCACCTGCGGCCCTCCACCTGCGGCCCTCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  865 GAGTACTGTGATGTGCCCTCTGCTCCACCTGCGGCCCTCCACCTGCGGCCCTCAG 924
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X07393
; DATABASE ENTRY DATE: 1995-03-27
; RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match             99.8%;   Score 1063.4;   DB 14;   Length 2509;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 1064;   Conservative  0;   Mismatches  0;   Indels  1;   Gaps  0;

Qy         1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       678 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 737

Qy        61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       738 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 797

Qy       121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       798 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 857

Qy       181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       858 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 917

Qy       241 GAGTACTGTGATGTGCCCTCCTGGTCTCCACCTGCGGGCCTGAGACAGTACAGCCAGCCTCAG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       918 GAGTACTGTGATGTGCCCTCCTGGTCTCCACCTGCGGGCCTGAGACAGTACAGCCAGCCTCAG 977

Qy       301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATGCTGCGCCGGTTCCTGTGCGGGGGCATACTCAGC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       978 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATGCTGCGCCGGTTCCTGTGCGGGGGCATACTCAGC 1037

Qy       361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCAGC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1038 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCAGC 1097

Qy       421 TCCTGCTGGATTCTCTGCCCGCCCACTGCTTCCGCCCCACCACTG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1098 TCCTGCTGGATTCTCTGCCCGCCCACTGCTTCCGCCCCACCACTG 1157

Qy       481 ACGGTGATCTTGGGCAGAACATACCGGGTGCCTGGCGAGGAGCAGAAATTTGAA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1158 ACGGTGATCTTGGGCAGAACATACCGGGTGCCTGGCGAGGAGCAGAAATTTGAA 1217

Qy       541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATTGCGG 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1218 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATTGCGG 1277

Qy       601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCACT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1278 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCACT 1337

Qy       661 GTGTGCCCTTCCCCCGGCGCCGGGACCTGGACTGGACCTCTCGGGC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1338 GTGTGCCCTTCCCCCGGCGCCGGGACCTGGACTGGACCTCTCGGGC 1397

Qy       721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGATGAAGGAGGCTCATGTC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1398 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGATGAAGGAGGCTCATGTC 1457

Qy       781 AGACTGTACCCATCCAGCCGCTGCACATCCAACACATTCACCGAC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1458 AGACTGTACCCATCCAGCCGCTGCACATCCAACACATTCACCGAC 1517

Qy       841 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCAGGCAAACTTGCACGACGCC 900
             ||||||||||||||||||||
```

```
Db      1518 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCAGGCAAACTTGCACGACGCC 1577

Qy       901 TGGCAGGGCGGATTCGGGAGGAGGCCGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTGGTG 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1578 TGGCAGGGCGGATTCGGGAGGAGGCCGGCCCCCTGGTGTGTCTGAACGATGGCCGATGACTTGGTG 1637

Qy       961 GGCATCATCAGCTGGGCGCTGGAGCTGTGACAGAGAGGATGTCCCGGTGCTGTACAAG 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1638 GGCATCATCAGCTGGGCGCTGGAGCTGTGACAGAGAGGATGTCCCGGTGCTGTACACCAAG 1697

Qy      1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGGTGA 1065
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      1698 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGGTGA 1742

RESULT 5
US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match             99.8%;   Score 1063.4;   DB 9;   Length 2519;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 1064;   Conservative  0;   Mismatches  0;   Indels  1;   Gaps  0;

Qy         1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       701 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 760

Qy        61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       761 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 820

Qy       121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       821 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 880

Qy       181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       881 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 940

Qy       241 GAGTACTGTGATGTGCCCTCCTGGTCTCCACCTGGGGGCCTGAGACAGTACAGCCAGCCTCAG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       941 GAGTACTGTGATGTGCCCTCCTGGTCTCCACCTGGGGGCCTGAGACAGTACAGCCAGCCTCAG 1000

Qy       301 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCGACATCGCCGGTGCCATC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1001 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCGACATCGCCGGTGCCATC 1060

Qy       361 TTTGCCAAGCACAGGAGGTCGCTTCTTCCAGGAGAGCTTCATCAGC 420
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      1061 TTTGCCAAGCACAGGAGGTCGCTTCTTCCAGGAGAGCTTCATCAGC 1120

Qy       421 TCCTGCTGGATTCTTCGCACATCACAACACATTAACGACAGTCACCTG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1121 TCCTGCTGGATTCTTCGCACATCACAACACATTAACGACAGTCACCTG 1180

Qy       481 ACGGTGATCTTGGGCAGAACATACCGGGTGTCCTGCGAGGAGGAGAAATTTGAA 540
             ||||||||||||||||||||
```

```
Qy      1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTGGTCAGCCTTACCGTGGCGAGCACGCAGAGC   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTGGTCAGCCTTACCGTGGCGAGCACGCACAGC  126

Qy     61 CTCACCGGATGCGGGTGCCTCCTGCCTCCTGCGTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 CTCACCGGAGCTGCGGGTGCCTCCTCCTGCGTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  186

Qy    121 TACCAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 TACCAGCACAGAACCCCAGTGCCAGCACTGGGCCTGGGCAAACATAATTACTGCCGG  246

Qy    181 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCAGGCTGACGTGG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 AATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCAGGCTGACGTGG  306

Qy    241 GAGTACTGTGATGTGCCCTCTGCCACCTGCGGGCCTGAGACAGTAGCCAGCCTCAG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 GAGTACTGTGATGTGCCCTCTGCCACCTGCGGGCCTGAGACAGTACAGCCAGCCTCAG  366

Qy    301 TTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCTGCAGCTGCCATC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 TTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCTGCAGCTGCCATC  426

Qy    361 TTTGCCAAGCACAGGAGTGCGCCGGACAGGCGGTTCGTGTGCGGGGGGCATACTCATCAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 TTTGCCAAGCACAGGAGTGCGCCGGACAGGCGGTTCGTGTGCGGGGGGCATACTCATCAGC  486

Qy    421 TCCTGCTGGATTCTCTCTGCCCGCCACTGCTTCCAGGAGAGTGTCCGCCGAGGAGGTTCGGAG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 TCCTGCTGGATTCTCTCTGCCCGCCACTGCTTCCAGGAGAGTGTCCGCCGAGGAGGTTCGGAG  546

Qy    481 ACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    547 ACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA  606

Qy    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  666

Qy    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGGGTGCCCAGGCTGGACGGAGTGAGCTCTCCGGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    667 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGGGTGCCCAGGCTGGACGGAGTGAGCTCTCCGGC  726

Qy    661 GTGTGCCTTCCCCCCGGCGGACCTGGTGTTGTGGACCGGAGTGTGAGCTGGACGGAGTGAGCTCTCCGGC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    727 GTGTGCCTTCCCCCCGGCGGACCTGCAGCTGCAGCCGGAGTGTGAGCTGGACGGAGTGAGCTCTCCGGC  786

Qy    721 TACGGCAAGCATGAGGCCTTGTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    787 TACGGCAAGCATGAGGCCTTGTCGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC  846

Qy    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    847 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  906

Qy    841 AACATGCTGTGTCTGGAGACACTCGGAGGCGGCGGCCCCAGGCAAACTTGCACGACGCC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    907 AACATGCTGTGTCTGGAGACACTCGGAGGCGGCGGCCCCAGGCAAACTTGCACGACGCC  966

Qy    901 TGCCAGGCGATTCGGGAGGCCCCCGTGTCTGTTCTGAACGATGGCCCGCCATGGCCCGATGACTTTGGTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    967 TGCCAGGCGATTCGGGAGGCCCCCGTGTCTGTTCTGAACGATGGCCCGCCATGGCCCGATGACTTTGGTG 1026

Qy    961 GGCATCATCAGCTGGGCGGGCCTGGGCGTGTGGACAGAAGGATGTGCCGGGGTGTGTACACAAAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GGCATCATCAGCTGGGCGGGCCTGGGCGTGTGGACAGAAGGATGTGCCGGGGTGTGTACACAAAG 1086

Qy   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1062
          |||||||||||||||||||||||||||||||||||||||||||
Db   1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
```

```
RESULT 6
US-09-987-455-2
; Sequence 2, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE:  2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
;   LENGTH: 1128
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: coding
;   OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match     99.7%;  Score 1062;  DB 11;  Length 1128;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

RESULT 7
US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match   99.7%; Score 1062; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGACTTCGGAATGGTCAGCGTACCGTGGCACGCACAGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 TCTGAGGGAAACAGTGACTGACTTCGGAATGGTCAGCGTACCGTGGCACGCACAGC 126

Qy   61 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127 CTCACCGAGTCGGGTGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 186

Qy 121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCTGGGCACTGGGCCTGGGCACTGCGCGG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 187 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCTGGGCACTGGGCCTGGGCACTGCGCGG 246

Qy 181 AATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 247 AATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGG 306

Qy 241 GAGTACTGTGATGTGCCCTCCGTCCACCTGCGGGCCTGAGACAGTCAGCCAGCCTCAG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 307 GAGTACTGTGATGTGCCCTCCGTCCACCTGCGGGCCTGAGACAGTCAGCCAGCCTCAG 366

Qy 301 TTTCGCATCAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 367 TTTCGCATCAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 426

Qy 361 TTTGCCAAGCACAAGGAGGAGGTCGCCCGAGGAGGTCGCCCGAGGAGGCCATACTCATCAGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 427 TTTGCCAAGCACAAGGAGGAGGTCGCCCGAGGAGGTCGCCCGAGGAGGCCATACTCATCAGC 486

Qy 421 TCCTGCTGGATTCTCTGCCGCCACCTGCTTCCGCCCCACCACCTG 480
         |||||||||||||||||||||||||||||||||||||||||||||
Db 487 TCCTGCTGGATTCTCTGCCGCCACCTGCTTCCGCCCCACCACCTG 546

Qy 481 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGGGGTGGTCCCTGGCAGAAATTTGAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 547 ACGGTGATCTTGGGCAGAACATACGGGGTGGTCCCTGGGGTGGTCCCTGGCAGAAATTTGAA 606

Qy 541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACAATGACATTGCG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 607 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACAATGACATTGCG 666

Qy 601 CTGCTGCAGCTGAAATGGATTCGTCCGCTGTCCGCTGTGCCCGGACTGACGCAGCGGTGCCGGCACT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 667 CTGCTGCAGCTGAAATGGATTCGTCCGCTGTCCGCTGTGCCCGGACTGACGCAGCGGTGCCGGCACT 726

Qy 661 GTGTGCCTCCCCGGCGACCTGCAGCTGCAGCGGACTGACGTGTGAGCTCTCCGGC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 727 GTGTGCCTCCCCGGCGACCTGCAGCTGCAGCGGACTGACGTGTGAGCTCTCCGGC 786

Qy 721 TACGGCAAGCATGAGGCATGAGGCCGGCTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 787 TACGGCAAGCATGAGGCATGAGGCCGGCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC 846

Qy 781 AGACTGTACCCATCCAGCCGCGTGCACATCACAACACATTACTTAACAGAACAGTCACCGAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 847 AGACTGTACCCATCCAGCCGCGTGCACATCACAACACATTACTTAACAGAACAGTCACCGAC 906

Qy 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGGGCCCCAGGCAAACTTGCAGACGCGC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 907 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGGGCCCCAGGCAAACTTGCAGACGCGC 966

Qy 901 TGCCAGGGCGATTCGGGAGCGCCCCCTGGTGTGTCTGACAGCGATGGCCGCATGCCGCCATGACTTTGGTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 967 TGCCAGGGCGATTCGGGAGCGCCCCCTGGTGTGTCTGACAGCGATGGCCGCATGCCGCCATGACTTTGGTG 1026

Qy 961 GGCATCATCAGCTGGGCCTGGGCCTGTGACGAAGAAGGATGTCCCGGGTGTGTACACAAAG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1027 GGCATCATCAGCTGGGCCTGGGCCTGTGACGAAGAAGGATGTCCCGGGTGTGTACACAAAG 1086

Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGGCACCG 1062
         ||||||||||||||||||||||||||||||||||||||||||
Db 1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGGCACCG 1128

RESULT 8
US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
;
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

Query Match   99.7%; Score 1061.8; DB 10; Length 2641;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1063; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGACTGACTTCGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 821 TCTGAGGGAAACAGTGACTGACTTCGGAATGGTCAGCCTACCGTGGCACGCACAGC 880

Qy   61 CTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 881 CTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 940

```
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
;
SEQ ID NO 433
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (82)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (159)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (264)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (278)
; OTHER INFORMATION: a, t, c or g
US-10-007-926A-433
```

```
Query Match          21.6%;     Score 230.2;  DB 12;   Length 329;
Best Local Similarity  97.5%;    Pred. No. 4.3e-63;
Matches    232;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy     828 AACAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCAGGCAA 887
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1 AACAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCCCCAGGNAA 60

Qy     888 CTTGCACGACGCCTGCCAGGGCGATTCGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCG 947
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      61 CTTGCACGACGCCTGCCAGGGNATTCGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCG 120

Qy     948 CATGACTTTGGTGGGCATCATCAGCTGGGCTGTGTGGACAGAAGGATGTCCCGGG 1007
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CATGACTTTGGTGGGCATCATCAGCTGGGCTGGGCTNTGGACAGAAGGATGTCCCGGG 180

Qy    1008 TGTGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGCGACCGTGA 1065
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGTGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGCGACCGTGA 238
```

```
RESULT 10
US-09-880-503-15
; Sequence 15, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
```

```
Qy     121 TACACAGCACGAACCCAGTGCCAGGCACTGGGCGCTGGGCAAACATAATTACTGCCGG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     941 TACACAGCACGAACCCCAGTGCCAGGCACTGGGCGCTGGGCAAACATAATTACTGCCGG 1000

Qy     181 AATCCTGATGGGATGCCAAGCCCTGTTGCCACGTGCCACGTGCCAGGCTGACGTGG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1001 AATCCTGATGGGATGCCAAGCCCTGTTGCCACGTGCCACGTGCCAGGCTGACGTGG 1060

Qy     241 GAGTACTGTGATGTGCCCCTCCTGTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1061 GAGTACTGTGATGTGCCCCTCCTGTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 1120

Qy     301 TTTCGGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1121 TTTCGGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 1180

Qy     361 TTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1181 TTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGC 1240

Qy     421 TCCTGCTGGATTCTCTCTGCCGCGCCACTGCTTCCGCCCCCCACCACCTG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1241 TCCTGCTGGATTCTCTCTGCCGCGCCACTGCTTCCGCCCCCCACCACCTG 1300

Qy     481 ACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1301 ACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAATTTGAA 1360

Qy     541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCCG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1361 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCCG 1420

Qy     601 CTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCGACT 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1421 CTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCGACT 1480

Qy     661 GTGTGCCTTCCCCGGCGGCAGCTGCAGCCTGCAGCTGGACGAGTGTGAGCTTCTCCGGC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1481 GTGTGCCTTCCCCGGCGGCAGCTGCAGCCTGCAGCTGGACGAGTGTGAGCTTCTCCGGC 1540

Qy     721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGCGGCTGAAGGAGGCTCATGTC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1541 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGCGGCTGAAGGAGGCTCATGTC 1600

Qy     781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTAACACAGACAGTCACCGAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1601 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTAACAGACAGTCACCGAC 1660

Qy     841 AACATGCTGTGTGTCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACTTGCACGACGCC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1661 AACATGCTGTGTGTCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACTTGCACGACGCC 1720

Qy     901 TGCCAGGGCGATTCGGAGGGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1721 TGCCAGGGCGATTCGGAGGGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1780

Qy     961 GGCATCATCAGCTGGGCGCTGGGGTGTGGACAGAAGATGTCCCGGGTGTGTACACAAAG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1781 GGCATCATCAGCTGGGCGCTGGGGTGTGGACAGAAGAAGATGTCCCGGGTGTGTACACAAAG 1840

Qy    1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1065
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1841 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCGTGA 1885
```

```
RESULT 9
US-10-007-926A-433
; Sequence 433, Application US/10007926A
; Publication No. US2003014353A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
```

```
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-880-503-15

Query Match     18.0%;  Score 191.6;  DB 10;  Length 1212;
Best Local Similarity   52.3%;  Pred. No. 1.5e-50;
Matches  549;  Conservative  0;  Mismatches  459;  Indels  42;  Gaps  4;
```

```
Qy   18 CTGCTACTTTGGAATGGGTCAGCCTACCGTGCACGCACCAGCCTCACCGGAGTCGGGTGC   77
         |||||||         ||||||
Db  147 CTGCTATGAGGGGAATGTGCTACTTTTACCGAGGAAAAGGCCAGCACTGACACCATGGGCCG  206

Qy   78 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACAGCAGCAGAACCC   137
Db  207 GCCCTGCCTGCGGACCTGCCACTGTCCTTCAGCAAACGTACCCATGCCCACAGATC   266

Qy  138 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGATGC   197
Db  267 TGATGCTCTTCAGCTGCGTGGGGCTGGGAAACATAATTACTGCTCGCAGGAACACCGGAG   326

Qy  198 CAAGCCCTGGTGCCACGTGCTGAAGAAGAACCGCAGGCCAGCGTGGGAGTACTGTGATGTGCC   257
Db  327 GCGAACCCTGGTGCTATGTGCAGGTGGGCCTAAGCGCCGTTGTCCAAGAGTGCATGGTGCA   386

Qy  258 CTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCA----------GCCTCAGTT   302
Db  387 TGACTGGCAGATGCAGATGGAAAATTAAAATTTCAGTGTGGCCAAAAGACTCTGAGGCCCCGCTT   446

Qy  303 TCGGCATCAAAGGAGGGGCTTCTTCCGCGACATGCGCTCCCACCCCTGGCAGGCTGCCATCTT   362
Db  447 TAAGATTATTGGGGAGAGAATTCACCACCATCGAACAACCAGCCCCTGGTTGCGGCCATCTA   506

Qy  363 TGGCCAAGCACAGAGAGCGTCGCCCCGAGAGGGGGTTCCTGTGCGGGGCATACTCATCAGCTC   422
Db  507 CAGGAGGCACCGGGGGGGGGCTCTGTCA---CCTACGTGTGTGAGGCAGCCTCATCAGCCC   563

Qy  423 CTGCTGGAATTCTCTCGCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCTGAC   482
Db  564 TTGCTGGGGTGATCAGCGCCACCACACTGGTTCATTGATTACCCAAAGAAGGAGGACTACAT   623

Qy  483 GGTGATCTTGGGCAGAACATACCGGGTGCCCTGGCGAGGAGGAGCAGAAATTTGAAGT   542
Db  624 CGTCTACCTGGGTCGGTGCTCAAGGCTTAACTCCAACACGCAAGGGAGATGAAGTTTGAGGT   683

Qy  543 CGAAAAATACATGTCATAAGGAATTCGATGATTGACA------CTTACGACAATGACAT   596
Db  684 GGAAAAACTTCATCCTCATCCTACACAGAGGACTACAGCGCTTGCTGACACGCTTGCCACCAACGACAT   743

Qy  597 TGCCCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGAGCAGCGTGGTCCG   656
Db  744 TGCCTTGCTGCGAAGATCCGTTCCAAGGAGGAGGCAGGTGTGCGCCAGCCATCCGGACTATACA   803

Qy  657 CACTGTGTGCCCTTCCCCGGCCGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTC   716
Db  804 GACCATCGCTCCTGCCCCTCCGATGTATAAGGATCCCCAGTTTGGCACAAGCTGTGTGAGATCAC   863

Qy  717 CGGCTACGGCAAGCAGCCATGAGGCGCGCATGAGGCGGCTGAAGGAGGCTCA   776
Db  864 TGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTATCCGGAGACTGAAAATGACTGT   923

Qy  777 TGTCAGACTTGTACCGCGCCGCTGCACATCCAGCCGCCCCACTACTACGGTCCTCTGAAGTCAC   836
Db  924 TGTCAAGCTGATTTCCCACCGGAGGTGTTCCAGCAGCCCCCACTACTACGGTCTCTGAAGTCAC   983

Qy  837 CGACAACATGCTGTGTGCTGGAGACACTCGGGACGCCCCCCAGGCAAACTTGCACGA   896
Db  984 CACCAAAATGCTATGTGCTATGTGCTGACCCCC------------AATGGAAAACAGA  1025
```

```
Qy  897 CGGCCTGCCAGGGCGAATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTT   956
Db 1026 TTCCTGCCAGGGAGACTCAGGGGAACCCCGTCGTCTGTTCCCTCCAAGGCCGCATGACTTT  1085

Qy  957 GGTGGGCATCATCAGCTGGGGCTGGGCTGGCTGGACAGAGAAGGATGTCCCGGGTGTGTACAC  1016
Db 1086 GACTGGAATTGTGAGCTGGCTGGGCCGTGGGCCCTGAAGGACAAGCCAGGCGTCTACAC  1145

Qy 1017 AAAGTTACCAACTACCTAGACTGGATTCG  1046
Db 1146 GAGAGTCTCCACACTTCTTACCCTGGATCCG  1175
```

```
RESULT 11
US-10-198-846-12748
; Sequence 12748, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION:  THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE:  2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 14084
; SEQ ID NO 12748
; LENGTH: 1415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 1414, 1415
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12748
```

```
Query Match     17.4%;  Score 185;  DB 14;  Length 1415;
Best Local Similarity   92.6%;  Pred. No. 2e-48;
Matches  249;  Conservative  0;  Mismatches  15;  Indels  5;  Gaps  5;
```

```
Qy    1 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGC   60
Db  906 TCTGAGGGAAAACAGTGACTGCTACTG-TACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAG-  963

Qy   61 CTCACCGAGTCGGGTGCCTCCT-GCCTCCCGTGGATCCTGAATTCCATGATCCTGATAGGCAAGGT  119
Db  964 CTCACCGAGTCGGGTCGGTGCCTCCTCCCGTGGATCCTGAATTCCATGATCCTGATAGGCAAGGT 1023

Qy  120 TTACACAGCACAGAACCCCAGTTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCG  179
Db 1024 TTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAAATTACTGCC 1083

Qy  180 GAATCCTGATGGGATGGCAAGACCCTCGTGCCACGTGCTGAA-GAACCGCAGGCCAGGCGACGT  238
Db 1084 GAAT-CTGATGGGATGGCAAGACCCTGTTTGCACGTGCTGAACGAGCCGCAGGCCAGGCGACGT 1142

Qy  239 GGGAGTACTGTGATGTGCCCTCTGCTCC  267
Db 1143 GGGAGTACTGTGATGTGGCCTCTGTCACC 1171
```

```
RESULT 12
US-09-735-705-122
; Sequence 122, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-705-122

Query Match    16.0%;   Score 170.8;  DB 9;  Length 1475;
Best Local Similarity   51.3%;   Pred. No. 7e-44;
Matches  551;   Conservative  0;  Mismatches 457;  Indels  66;  Gaps  4;

Qy   18 CTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACCTCACCGAGTCGGGTGC  77
         ||||||   ||||||||||
Db  287 CTGCTATGAGGGAAATGGTCACTTTTACCGAGGAAAAGGCCAGCACTGACACCAGGACCG 346

Qy   78 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATGAGCTGACCACTGA  137
         |||||| | ||   ||| || |||  |||
Db  347 GCCCTGCCTCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCCCACAGATC 406

Qy  138 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 197
         ||| ||||| |||||||||  |||||||||||||||
Db  407 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCGAGGAACCCAGACAACCGGAG 466

Qy  198 CAAGCCCTGGTGCCACGTGCTGAAGAACCAGCTGGGACTGGGACTGTGTGATGTGCC 257
         |||| ||||||||   ||| |||
Db  467 GCGACCCTGGTGCTATGTGCAGTGTGGGCCTAAAGCCGCTTGTCCAAGAGTGCATGGGTGCA 526

Qy  258 CTCCTGCTCCACCTGCGGCCTGAG------------------------ 281
         |||||||  ||||
Db  527 TGACTGCGCAGATGGAAAAAAGCCCTCTCTCCTCCAGAGAATTAAAATTTCAGTGTGG 586

Qy  282 ---ACAGTACAGCCAGCCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTC 338
            | || ||  ||| || ||  || |||  ||||||  ||||||
Db  587 CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGAGAA 646

Qy  339 CCACCCCTGGCAGCTGCCATCTTCTGCCAAGCACAGAGAGGTCGCCCGAGAGCGGTTCCT 398
         |||||  |||| |||||  ||  |||||||||||   |||
Db  647 CCAGCCCTGGTTTGCGGCCATCTACAGGAGGCACCGGGG--GGGCTCTGTGTCACCTACGT 703

Qy  399 GTGCGGGGGCATACTCATTCATCAGCTCCTGGATTCTCTGCCGGCCCACTGCTTCCAGGA 458
         || ||| ||   |||| || |||||||| |||||| |||
Db  704 GTGTGGAGCAGCCTCATCAGCCCCTTGCTGGGTGATCAGCGCCACCACATCGCTTCATTGA 763

Qy  459 GAGGTTTCCGCCCACCACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGCTGCCTGG 518
         |||| |||||||||||||||||
Db  764 TTACCCAAAGAAGGAGGACTACATCGTCTACCTGGGTCGCTCAAGGCTTAACTCCAACAC 823

Qy  519 CGAGGAGGACAGAAATTTGAAGTGCAAAAAATACATTGTCCATAAGGAATTCGATGATGA 578
         ||| |||||||  ||| |||||   || || |||
Db  824 GCAAGGGGGAGATGAAGTTTGAGGTGGAAAAACCTACACAAAGGACTACACAGCGCTGA 883

Qy  579 CA------CTTACGACAATGACATTGCCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTG 632
         ||      ||  ||||  |||||
Db  884 CACGGCTTGCTCCACCACAACGACATTGCCTTGCTTGCTGAAGATCCGGTTCCAAGGAGGAGGCAGGTG 943
```

```
Qy  633 TGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCGGCGGCGACCTGCAGCTGCC 692
         |||||    ||||   ||  ||||
Db  944 TGGCAGCCATCCGGACTATACAGACCATCTGCCTGCCCTGCGATGTATAACGATCCCCA 1003

Qy  693 GGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTA 752
         |||||||||| |||||  || |  ||
Db 1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy  753 TTCGGAGGGCTGAAGAGGCTCATGTCAGATGTGCAGACTGTACCCATCCAGCCGCTGCACATCACA 812
         | |||||| | |||||
Db 1064 TCCGGAGAGCTGAAGATGATGACTGTTGTGAACAGTGAAGCTGAATTCCCACCGGGAGTGTCAGCAGCC 1123

Qy  813 ACATTACTTAACAGAACAGTCACCGCCAACAGTGCTGTGTGCGGAGACACTCGGAGCGGG 872
          | ||||  ||  ||||||
Db 1124 CCACTACACGGCTCTGAAGTCACCACCAAAATGCTGTGTGCTGTGACCACCAGTG---- 1180

Qy  873 CGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGCAATTCGGGAGGCCCCCTGGTGTG 932
         |   |||||||||  |||| ||||||||  ||
Db 1181 ----------GAAAACAGATTCCTGCCAGGGACTCAGGGGGACCCCCTCGTCTG 1225

Qy  933 TCTGAACGATGGCGCATGACTTTGGTGGCATCATCATCATCAGCTGGGCCTGGGCGTGTGGACA 992
         |||||  |||||||||   ||||||
Db 1226 TTCCCTCCAAGGCCGATGACTTTGACTGGAATTGTGACTGGGTGGGCCGTGGGATGGTGCCCT 1285

Qy  993 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACACCTAGACTGGATTCG 1046
         ||||  ||
Db 1286 GAAGGACAAGCCAGGCGTCTACACGACGAGTCTCACACTTCTTACCTGGATCCG 1339
```

```
RESULT 13
US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US2002011513.9A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-850-716A-122

Query Match    16.0%;   Score 170.8;  DB 10;  Length 1475;
Best Local Similarity   51.3%;   Pred. No. 7e-44;
Matches  551;   Conservative  0;  Mismatches 457;  Indels  66;  Gaps  4;

Qy   18 CTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACCTCACCGAGTCGGGTGC  77
         ||||||   ||||||||||
Db  287 CTGCTATGAGGGAAATGGTCACTTTTACCGAGGAAAAGGCCAGCACTGACACCAGGACCG 346

Qy   78 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATGAGCTGACCACTGA  137
         |||||| | ||   ||| || |||  |||
Db  347 GCCCTGCCTCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCCCACAGATC 406

Qy  138 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 197
         ||| ||||| |||||||||  |||||||||||||||
Db  407 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCGAGGAACCCAGACAACCGGAG 466

Qy  198 CAAGCCCTGGTGCCACGTGCTGAAGAACCAGCTGGGACTGGGACTGTGTGATGTGCC 257
         |||| ||||||||   ||| |||
Db  467 GCGACCCTGGTGCTATGTGCAGTGTGGGCCTAAAGCCGCTTGTCCAAGAGTGCATGGGTGCA 526

Qy  258 CTCCTGCTCCACCTGCGGCCTGAG------------------------ 281
         |||||||  ||||
Db  527 TGACTGCGCAGATGGAAAAAAGCCCTCTCTCCTCCAGAGAATTAAAATTTCAGTGTGG 586
```

```
;  NUMBER OF SEQ ID NOS: 467
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 122
;  LENGTH: 1475
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-897-778-122

  Query Match             16.0%;   Score 170.8;  DB 10;  Length 1475;
  Best Local Similarity   51.3%;   Pred. No. 7e-44;
  Matches  551; Conservative   0;  Mismatches 457;  Indels  66;  Gaps   4;

Qy          18 CTGCTACTTGGGAATGGTCAGCTCAGGTCGGACGCACCGTGGCACAGCCTCACCGAGTCGGGTGC    77
               |||||       ||||||   ||||||  | ||    |||||  || ||||
Db         287 CTGCTATGAGGGGAATGGTCACTTTTACCGGAGAAAGCCAGCACTGACACCATGGGCCG    346

Qy          78 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACGACAGAACCC    137
               |||||| |||| ||| |||||||  ||  | |
Db         347 GCCCTGCCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCCACAGATC    406

Qy         138 CAGTGCCCAGGCACTGGGCTGGCAAACATAATTACTGCCGAAATCTGATGGGGATGC    197
               |||||||||||  ||||||  || || |||| ||| ||||   |||
Db         407 TGATGCTCTTCAGCTGGGGCTGGGGAAACATAATTACTGCAGGAAACCCAGAACAACCGAG    466

Qy         198 CAAGCCCTGGTGCCACTGTTCCGCACTGTGCCTTCCCCGGCGACCTGCAGCTGCC    257
               ||  |||||||| || || || ||   ||  || |||  || ||||||| | || 
Db         467 GCGACCCTGGTGCTATGTGCAGTGGGCCTAAAGCCGGTTGTCCAAGAGTGCATGGTGCA    526

Qy         258 CTCCTGCTCCACCTGCGGCCTGAG----------------------------    281
               |||||||||    | |
Db         527 TGACTGCCGAGATGGGAAAAACGCCCTCTCCTCCTCCAGAAGAATTAAAATTCAGTGTGG    586

Qy         282 --ACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTC    338
                  ||||||||| ||| ||| |||    ||||||| |||| || || | ||||
Db         587 CCAAAGACTCTGAGGCCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGAGAA    646

Qy         339 CCACCCCTGGCAGGCTGCCATCTTGCGGCCATCACAGAGGAGGCACCGGGG---GGGCTCGTGTCACCTACGT    398
               |||||||||| ||||| |||| |||   ||||  ||| |||||||   ||||||   ||||
Db         647 CCAGGCCCTGGTTTGCGGCCATCATCACAGGAGGCACCGGGG---GGGCTCGTGTCACCTACGT    703

Qy         399 GTGCGGGGGGCATACTCATCAGCTCCTGCGGATTCTCTGCCGACCCACGTGCTTCCAGGA    458
               |||||        | || ||| ||  ||| |  | ||| | |||| |||  ||  ||||
Db         704 GTGTGGAGGCAGCAGCCTCATCAGCTGGGTGATCAGCGGCCGCCACACACTGCTTCATTGA    763

Qy         459 GAGGTTTCCGCCCACCACTGAGCTCTGTGAGCTTGGGCATGTGGGGCCTGGGCTGTCCCTGG    518
               |||| |||||| | || |||  |||| |||  |||||  ||||| |||  ||||  || |||
Db         764 TTACCAAAGAAGGAGGACTACATCGTCTACCTGGGTCGCTCAAGGTTAACTCCAACAC    823

Qy         519 CGAGGAGGCAGCAGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGA    578
               ||||||||||  || |||||||  | |||   ||   |||  ||||||  ||
Db         824 GCAAGGGAGATGAAGTTTGAGGTGGAAAAACCTCATCCTACACAAGGACTACAGCGCTGA    883

Qy         579 CA------CTTACGACAATGACATTGCGCTGCTGCAGCTGAATCGGCTGGTCATGGTG    632
               ||       ||||  |||| |  ||| ||||||||  | ||||  || ||||| ||||
Db         884 CACGCTTGCTCTCACCACACAGACGACGATGCGTCCGTGCTGCGCAGCTGGTGCAGGTG    943

Qy         633 TGCCCAGGAGGACGGCGTGGCGGAGTGGAGCTCTCCGGTCCGGCAAGATCCGTGGTCCCTGG    692
               ||||| |||||| |||||||||||||||||||| |||||  | ||  ||  ||| ||||||
Db         944 TGGCGGCATCCGTGGTGCTGTGTGCTGGGTCGCTCACCCGCGCCAGCTGCCCAGCTGCC    1003

Qy         693 GGACTGGACGGAGGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCGGTTCCT    752
               ||||| || ||||||||| || ||||| |  || ||| ||||| ||| |||
Db        1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA    1063

Qy         753 TTCGGAGCGGCTGAAGAGGCTTCATGTGCAGACTGTTCAGACTGTACGGCTGCACATCACA    812
               ||||| ||||  || |||||||  ||||| || ||| | ||| | | |  || || |||
Db        1064 TCCGGAGCAGCTGAAGATGAGTTGTTGTGGTGGAAAACCTCATCCTACACAAGGACTACAGCGCC    1123

Qy         813 ACATTACTTAACAGACAGTACACCGACCAACATGCTGTCTGCGGACACTCGGAGCGG    872
               |||||            | ||  | || ||||   ||
Db        1124 CCACTACACGGCTCTGAAGTCACCAAAATGTGTGTCTGCTGACCCACAGTG---    1180

Qy         873 CGGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGGCCGATTCGGGAGGGGCCCCTGGTGTG    932
```

```
Db  1181 ------------GAAAACAGATTCCTGCCAGGGAGACTCAGGGGGACCCCTCGTCTG 1225
              ||   |||||||||| || ||  || || |||| ||||| || ||
Qy   933 TCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACA 992
         ||||||||||||||||| |||| ||   ||   |||||  || |||||||| |||
Db  1226 TTCCCTCCAAGGCCGCATGACTTTGACTGAATTGTGAGCTGGGGCCGTGGGATGTGCCCT 1285

Qy   993 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTCGGATTCG 1046
         ||||||   || |||| |   |||||||| ||||||||||||||||| ||
Db  1286 GAAGGACAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCCTGGATCCG 1339

RESULT 15
US-09-466-396A-122
; Sequence 122, Application US/09466396A
; Publication No. US2003011976A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-122

Query Match          16.0%;   Score 170.8;   DB 11;   Length 1475;
Best Local Similarity 51.3%;  Pred. No. 7e-44;
Matches  551;  Conservative  0;  Mismatches 457;  Indels 66;  Gaps 4;

Qy    18 CTGCTACTTTGGGAATGGGTCAGCTACCGTGGCACGCACAGCCTCACCGAGTCGGGTGC 77
         ||||||        ||||||            ||||   ||||| ||      ||
Db   287 CTGCTATGGAGGGGAATGGTCACTTTTACCGAGGAAAAGGCCAGCACTGACCATGGGCCG 346

Qy    78 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACAGCACAGAACCC 137
         ||||||||             |||||   |   ||     |||||||||
Db   347 GCCCTGCCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCATGCCCACAGATC 406

Qy   138 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGATGC 197
         || |      ||||||||           |||||| ||| ||   ||   |
Db   407 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCAGGAACCCAGAACCGGAG 466

Qy   198 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCTGACGTGGGAGTACTGTGATGTGCC 257
         ||||||||              ||||||||       |||||  ||||||||||
Db   467 GCGACCCTGGTGCTGCTATGTGCAGGTGGGCCTAAAGCCGCTTGTCCAAGAGTGCATGGTGCA 526

Qy   258 CTCCTGCTCCACCTGCGGCCTGAG------------------------ 281
         ||||||||||||||||            
Db   527 TGACTGCGCAGATGGAAAAAGCCCTCTCCTCTCCAGAAGAATTAAAATTTCAGTGTGG 586

Qy   282 ---ACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTC 338
               |||||     ||||| |||||   |||     |||||  |||||||| |||
Db   587 CCAAAAGACTCTGAGGGCCCCGCCTTTAAGGATTATTGGGGAGAATTCACCACCATCGAGAA 646

Qy   339 CCACCCCTGGCAGCTGCCATCTTGCCAAGCACAGAGGTCGCCCCGAGAGCGGTTCCT 398
         || |||||| ||||||||||||          || ||||| ||||||||||||
Db   647 CCAGCCCCTGGTTGCGGCCATCTACAGGAGGCACCGGGGG--GGGCTCTGTCACCTACGT 703

Qy   399 GTGCGGGGGCATACTCATCAGCTGGAATCTTCTGCGCCGCCACTGCTTCCAGGA 458
         || ||   ||||| |||||||||||   |||||||||   |||| ||| |||||
Db   704 GTGTGGAGGCAGCCTCATCAGCCCCTTGCTGGGTGATCAGCAGCCACACACTGCTTCATTGA 763

Qy   459 GAGGTTTCCGCCCACCACCTGACGGTGATCATTGGGCAGAACAATACCGGGTGTCCCTGG 518
         |||||| |||||  |||||||||||  |||| |||||||||  |||  ||||||| |||
Db   764 TTACCCAAGAAGAGGACGACTACATCGTCTACCTGGGTCGCCTCAAGGCTTAACTCCAACAC 823

Qy   519 CGAGGAGCAGCAGAATTTGAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGA 578
         || ||   ||||| |||||||||||        ||| |  ||| |||||||| || |||
```

Db   824 GCAAGGGGAGATGAAGTTTGAGGTGGAAAACCTCATCCTACACAAGGACTACAGCGCTGA 883
              || |||||| |||||| || |||| ||| ||
Qy   579 CA------CTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCTG 632
         || |||| ||| || || || ||||| |||| ||||||||
Db   884 CACGCCTTGCTCACCACACGACATTGCCTTGCTGAAGATCCGTTCCAAGGAGGGCAGGTG 943

Qy   633 TGCCCAGGAGAGCAGCGGTGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGAGCTGCC 692
         | || ||| ||||||||| ||||||||| || |||| |||| ||| |
Db   944 TGCCAGCCATCCGGACTATACAGACCATCTGCTGCCCTGCCTGATGTATAACGATCCCCA 1003

Qy   693 GGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTA 752
         ||||||| ||||| | ||||||| || |||| |||| || |||||||||
Db 1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy   753 TTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACA 812
         ||||||| ||||| ||| ||||| || || | | ||| || ||||
Db 1064 TCCCGAGCAGCTGAAGATGACTGTTGTGAAGCTGAATTCCCACCGGAGTGTCAGCAGCC 1123

Qy   813 ACATTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGG 872
         || | ||||||| ||| |||| | |||||||| ||||||
Db 1124 CCACTACTACGGCTTCTGAAGTCACCACCAAAATGCTGTGCTGTGCTGACCACAGTG--- 1180

Qy   873 CGGGCCCCAGGCAAACTTGCCACGAGCGCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTG 932
         | ||||| ||| |||||| ||| || | |
Db 1181 --------GAAAACAGATTCCTGCCAGGGAGACTCAGGGGGACCCCTCGTCTG 1225

Qy   933 TCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACA 992
         ||||||||||||||||| |||| || || ||||| || |||||||| |||
Db 1226 TTCCCTCCAAGGCCGCATGACTTTGACTGAATTGTGAGCTGGGGCCGTGGGATGTGCCCT 1285

Qy   993 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTCGGATTCG 1046
         |||||| || |||| | |||||||| ||||||||||||||||| ||
Db 1286 GAAGGACAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCCTGGATCCG 1339

Search completed: August 19, 2003, 14:22:25
Job time : 302.529 secs

OM nucleic - nucleic search, using sw model

Copyright (c) GenCore version 5.1.6 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 22:44:33 ; Search time 3115.64 Seconds
(without alignments)
8307.845 Million cell updates/sec

Title: US-09-987-455-4
Perfect score: 1065
Sequence: 1 tctgaggaaacagtgactg.........gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215223805 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
 1: em_estba:*
 2: em_esthum:*
 3: em_estin:*
 4: em_estmu:*
 5: em_estov:*
 6: em_estpl:*
 7: em_estro:*
 8: em_htc:*
 9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%Query Match	Length	DB	ID	ID	Description
c 1	895	84.0	930	13	BX329047	BX329047	BX329047
2	853	80.1	912	13	BU149958	BU149958	AGENCOURT
3	850.6	79.9	962	13	BX389609	BX389609	BX389609
4	818.6	76.9	891	13	BU146167	BU146167	AGENCOURT
5	815.8	76.6	993	13	BQ278177	BQ278177	AGENCOURT
6	811.2	76.2	883	13	BQ690978	BQ690978	AGENCOURT
7	807.6	75.8	918	13	BX389608	BX389608	BX389608
8	796.2	74.8	930	13	BU845014	BU845014	AGENCOURT
9	782.8	73.5	888	13	BU157720	BU157720	AGENCOURT
10	767.4	72.1	916	12	BI765247	BI765247	603050338
11	765.2	71.8	950	13	BQ927899	BQ927899	AGENCOURT
12	706.8	66.4	948	10	BB616613	BB616613	601278864
13	697.4	65.5	958	12	BG763582	BG763582	602735987
14	694.4	65.2	974	13	BQ960216	BQ960216	AGENCOURT
15	693.4	65.1	709	10	BE732704	BE732704	601571366
16	685.2	64.3	1060	12	BM56999	BM56999	AGENCOURT
17	677	63.6	1042	10	BE732413	BE732413	60159644
18	672	63.1	957	13	BQ687538	BQ687538	AGENCOURT
19	668.8	62.8	948	13	BQ687779	BQ687779	AGENCOURT
20	665.2	62.5	704	9	AU134301	AU134301	AU134301
21	658.2	61.8	903	13	BQ689095	BQ689095	AGENCOURT
22	656.2	61.6	830	13	BQ896938	BQ896938	AGENCOURT
23	655.2	61.5	904	13	BQ689840	BQ689840	AGENCOURT
24	654.6	61.5	911	13	BQ684734	BQ684734	AGENCOURT
25	654.6	61.5	867	13	BU179903	BU179903	AGENCOURT
26	652.4	61.3	818	9	AU124602	AU124602	AU124602
27	651	61.1	818	12	BM006474	BM006474	603615409
28	649.6	61.0	797	13	BQ879911	BQ879911	AGENCOURT
29	642.2	60.3	1037	13	BU157493	BU157493	AGENCOURT
30	631.2	59.3	813	13	BG120780	BG120780	602348540
31	629.8	59.1	663	9	AU134463	AU134463	AU134463
32	624	58.6	624	10	BE549489	BE549489	601278502
33	585.4	55.0	587	10	BE394944	BE394944	601308225
34	561	52.7	561	9	AL047284	AL047284	DKFZp586H
35	553.4	52.0	555	14	CB128855	CB128855	K-EST0178
36	535.4	50.3	537	14	CB155052	CB155052	K-EST0213
37	531	49.9	964	12	BG769384	BG769384	602742678
38	517.4	48.6	905	13	BU859271	BU859271	AGENCOURT
39	514.8	48.3	883	14	CD359959	CD359959	AGENCOURT
40	514.4	48.3	516	14	CB140923	CB140923	K-EST0194
41	496.4	46.6	916	13	BU157241	BU157241	AGENCOURT
42	487.8	45.8	527	9	AW630119	AW630119	hh78e06.y
43	485.8	45.6	553	9	AW394127	AW394127	MR2-TT001
44	477.8	44.9	917	13	BU190904	BU190904	AGENCOURT
45	475.8	44.7	937	13	BQ679501	BQ679501	AGENCOURT

ALIGNMENTS

RESULT 1
BX329047/c
LOCUS   BX329047   930 bp   mRNA   linear   EST 01-MAY-2003
DEFINITION BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 3-PRIME, mRNA sequence.
ACCESSION   BX329047
VERSION   BX329047.1 GI:30307809
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4498.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAI0212D08_C
            S01980 1&cluster=4498.r. Contact : Feng Liang Email :
            fliang@lifetech.com URL : http://fulllength.invitrogen.com/
            InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID :

FEATURES
　source
　　　　CS0BAIO21ZD08_CS01980.1
　　　　Location/Qualifiers
　　　　1..930
　　　　/organism="Homo sapiens"
　　　　/mol_type="mRNA"
　　　　/db_xref="taxon:9606"
　　　　/clone="CS0DI039YK15"
　　　　/tissue_type="PLACENTA COT 25-NORMALIZED"
　　　　/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
　　　　/note="1st strand cDNA was primed with a NotI-oligo(dT)
　　　　primer. Five prime end enriched, double-strand cDNA was
　　　　digested with Not I and cloned into the Not I and EcoR V
　　　　sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT　　184 a　　268 c　　282 g　　194 t　　2 others
ORIGIN

Query Match　　912; Best Local Similarity　84.0%; Score 895; DB 13; Length 930;
Best Local Similarity　98.6%; Pred. No. 5.7e-223;
Matches 912; Conservative　0; Mismatches 12; Indels 1; Gaps 1;

Qy　　　1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60
　　　　　|||||||||||||||||||||||||||||| || ||||||| ||||||||||||||||||
Db　　 927 TTTGAGGGAAACAGTGACTGCTAC-TTGGGAATGGTCAGCCTACGTGGGCACGCACAGC 869

Qy　　　61 CTCACCGAGTCGGGTGCCTCCTCCGCTCCCTGGAATTCCATGATCCTGATAGGCAAGGTT 120
　　　　　|||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db　　 868 CTCACCGAGTCGGGTGCTNCCTGCCTCCCTGGAATTCCATGATCCTGATAGGCAAGGTT 809

Qy　　 121 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 180
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 808 TACACAGCACAGAACNCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG 749

Qy　　 181 AATCCTGATGGGATGCCAAGCCCTGGTTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 748 AATCCTGATGGGATGCCAAGCCCTGGTTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 689

Qy　　 241 GAGTACTGTGATGTGCCCTCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAGCCTCAG 300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 688 GAGTACTGTGATGTGCCCTCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAGCCTCAG 629

Qy　　 301 TTTCGCATCAAAGAGGAGGGCTCTTGCCGACATCGCCTCCACCCTGGCAGGCTGCCATC 360
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 628 TTTCGCATCAAAGAGGAGGGCTCTTCGCCGACATCGCCTCCACCCTGGCAGGCTGCCATC 569

Qy　　 361 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTCGCGGGGGCATACTCATCAGC 420
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 568 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTCGCGGGGGCATACTCATCAGC 509

Qy　　 421 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 480
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 508 TCCTGCTGGATTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG 449

Qy　　 481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA 540
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 448 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAA 389

Qy　　 541 GTCGAAAAATACATTGTCCATAGGAATTCGATGATGAATCGACAATGACATTGCG 600
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 388 GTCGAAAAATACATTGTCCATAGGAATTCGATGATGAATCGACAATGACATTGCG 329

Qy　　 601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCGGTGCCCAGGAGGAGCAGCGTGGTCCGGACT 660
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 328 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCGGTGCCCAGGAGGAGCAGCGTGGTCCGGACT 269

Qy　　 661 GTGTGCGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGACCGGAGTGTGAGCTTCCGGC 720
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 268 GTGTGCGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGACCGGAGTGTGAGCTTCCGGC 209

Qy　　 721 TACGGCAAGCATGAGGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 208 TACGGCAAGCATGAGGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 149

Qy　　 781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTTAACAGACAGTCACCGAC 840

```
AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
COMMENT    Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen. This sequence belongs to sequence cluster
           4498.r For more information about this cluster, see
           http://www.genoscope.cns.fr/
           cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_3&cluster=4498.r.
           Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_3.

FEATURES             Location/Qualifiers
     source          1..962
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI039YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      203 a    290 c    277 g    191 t     1 others
ORIGIN

Query Match             79.9%;  Score 850.6;  DB 13;  Length 962;
Best Local Similarity   96.4%;  Pred. No.: 2.3e-211;
Matches    891;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    2;

Qy    142 GCCCAGGCACTGGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGCCAAG 201
              |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db      1 GCCCAGGCACTGGNGGCCTGGGCAAACTAATTACTGCCGGAATCCTGATGGGGATGCCAAG  60

Qy    202 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGGTGCCCCTCC 261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGGTGCCCCTCC 120

Qy    262 TGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTCCGCATCAAAGAGAGGGCTC 321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTCCGCATCAAAGAGAGGGCTC 180

Qy    322 TTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCG 381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCG 240

Qy    382 CCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCAGCTCATCAGCTCCTGCGATTCTCTGCC 441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCAGCTCATCAGCTCCTGCGATTCTCTGCC 300

Qy    442 GCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACA 501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACA 360

Qy    502 TACCGGGGTGGTCCCTGGCCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCAT 561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TACCGGGGTGGTCCCTGGCCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCAT 420

Qy    562 AAGGAATTCGATGATGACTTACGACAATGACATTGCCGCTGCAGCTGAAATCGAT 621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAGGAATTCGATGATGACTTACGACAATGACATTGCCGCTGCAGCTGAAATCGAT 480

Qy    622 TCGTCCCGCTGCCCAGGAGAGCAGCAGCGGTGTCCCGCACTGTGCCTTCCCCGCGGAC 681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCGTCCCGCTGCCCAGGAGAGCAGCAGCGGTGTCCCGCACTGTGCCTTCCCCGCGGAC 540

Qy    682 CTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGCCAAGCATGAGGCCTTG 741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGCCAAGCATGAGGCCTTG 600
```

```
Db     61 TGGGGATGCCAAGCCCTGGTGCCACGTGCTGCCACGCTGGGAGTACTACTG 120

Qy    249 TGATGTGCCCTCTGCTCCACCTGCGGCGGCCTGAGACAGTACAGCCAGTTCCGCAT 308
Db    121 TGATGTGCCCTCCTGCTCCACCTGCGGCGGCCTGAGACAGTACAGCCAGTTCCGCAT 180

Qy    309 CAAAGGAGGAGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGGCATCTTTGCCAA 368
Db    181 CAAAGGAGGAGTCGCCCGAGAGCCGGTTCCTGTGCGGGGGCATACTCAGCTCCTGCTG 240

Qy    369 GCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCATCAGCTCCTGCTG 428
Db    241 GCACAGGAGTCGCCCGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCATCAGCTCCTGCTG 300

Qy    429 GATTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACCTGACGGTGAT 488
Db    301 GATTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCACCACCTGACGGTGAT 360

Qy    489 CTTGGGCAGAACATACCGGGGTGTCCCTGGCCGGAGGAGGAGCAGAAATTTGAAGTCGAAAA 548
Db    361 CTTGGGCAGAACATACCGGGGTGTCCCTGGCCGGAGGAGGAGCAGAAATTTGAAGTCGAAAA 420

Qy    549 ATACATTGTCCATAAGGAATTCGATGATGACTTACGACAATGACATTGCCGCTGCTGCA 608
Db    421 ATACATTGTCCATAAGGAATTCGATGATGACTTACGACAATGACATTGCCGCTGCTGCA 480

Qy    609 GCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCAGCGTGTCCCGCACTGTGCCT 668
Db    481 GCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCAGCGTGTCCCGCACTGTGCCT 540

Qy    669 TCCCCGCGGAGCCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAA 728
Db    541 TCCCCGCGGAGCCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAA 600

Qy    729 GCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTCGAAGGAGGCTCATGTCGACTGTA 788
Db    601 GCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTCGAAGGAGGCTCATGTCGACTGTA 660

Qy    789 CCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGACAACATGCT 848
Db    661 CCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGACAACATGCT 720

Qy    849 GTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGG 908
Db    721 GTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGG 780

Qy    909 -CGATTCGGGAG-CCCCCTGGTGTCTGAACGATGGCCGATGACTTTGGTGGGGCATC 966
Db    781 NCGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGATGACTTTGGTGGGGCATC 840

Qy    967 ATCAGCTGGGGCGTTGGGGCCTGGGGCTGTGACAGAAGGATGT-CCCGGGTGTGTACAAAGGTTAC 1025
Db    841 ATCAGCTGGGGCGTTGGGGCCTGGGGCTGTGACAGAAGGATGTCCCCGGGTGTGTACAAAGGTTAC 900

Qy   1026 CAACTACCTAG 1036
Db    901 AAACTACCTAG 911
```

RESULT 3
BX389609

```
LOCUS       BX389609                 962 bp    mRNA    linear    EST 08-MAY-2003
DEFINITION  BX389609 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389609
VERSION     BX389609.1  GI:30463470
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 962)
```

```
Qy   742 TCTCCTTTCTATTCGGAGCGGCTGAAGAGGAGGCTCATGTCAGACTGTACCCATCCAGCGGC  801
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTCCTTTCTATTCGGAGCGGCTGAAGAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGC  660

Qy   802 TGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTCGGAGAC      861
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTCGGAGAC      720

Qy   862 ACTCGGAGCGGCGGCGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGGCGATTCGGGAGGC  921
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCGGAGCGGCGGCGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGGCGATTCGGGAGGC  780

Qy   922 CCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGGGCCTG     981
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCCTGGGCCTG     840

Qy   982 GGCTGCTGGACAGAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGG  1041
         ||||||||  |||||     |||||||  |||   |||
Db   841 GGCTGCTGGAACAGAAGGATGTCCCGGTGTCCCGGTGTTGTTCCCACCAAG--TTACCACTTACTAAACTGG  898

Qy  1042 ATTCGTGACAACATGCGACCGTGA  1065
         |||||  ||  ||||||||||||
Db   899 ATTCGGGAC-ACATGCGACCGGGA   921
```

```
RESULT 4
BU146167
LOCUS       BU146167                 891 bp    mRNA    linear    EST 03-SEP-2002
DEFINITION  AGENCOURT_7981262 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6185661 5', mRNA sequence.
ACCESSION   BU146167
VERSION     BU146167.1  GI:22659699
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 891)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13576 row: m  column: 22
            High quality sequence stop: 665.
FEATURES             Location/Qualifiers
     source          1..891
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6185661"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site 2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGGCGAGCCGCGGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."

BASE COUNT      178 a    266 c    271 g    176 t
ORIGIN
```

```
Query Match            76.9%;
Best Local Similarity  98.2%;     Score 818.6;  DB 13;  Length 891;
Matches  849;                     Pred. No. 5.2e-203;
Conservative  0;    Mismatches  14;    Indels  2;    Gaps  2;
```

```
Qy   136 CCCAGTGCCCAGGCACTGGGCTCGGCCAAACATAATTACTGCGAAACATAATTACTGATGGGAT  195
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCCAGTGCCCAGGCACTGGGCTCGGCCAAACATAATTACTGCGAAACATAATTACTGATGGGAT   60

Qy   196 GCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGCTGACGTGGGAGTACTGTGATGTG      255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGCTGACGTGGGAGTACTGTGATGTG      120

Qy   256 CCCTCCTGCTCCACCTGGGCTGAGACAGTACAGCCAGCCTCAGTTCGCATCAAAGGA      315
         |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGCCAGCCTCAGTTCGCATCAAAGGA     180

Qy   316 GGGCTCTTCGCCGACATGCCTCCACCCCTGCCAGGCTGCCATCTTTGCCAAGCACAGG      375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGGCTCTTCGCCGACATGCCTCCACCCCTGCCAGGCTGCCATCTTTGCCAAGCACAGG      240

Qy   376 AGGTCGCCCGGAGAGCGGTTCCTGTGCGTGCGGGGCATACTCATCAGTCCTGCTGGATTCTC  435
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AGGTCGCCCGGAGAGCGGTTCCTGTGCGTGCGGGGCATACTCATCAGTCCTGCTGGATTCTC  300

Qy   436 TCTGCCGCCACTGCTTCCCAGGAGGAGTTTCCGCCCCACCACTGACGGTGATCTTGGGC     495
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCTGCCGCCACTGCTTCCCAGGAGGAGTTTCCGCCCCACCACTGACGGTGATCTTGGGC     360

Qy   496 AGAACATACCGGTGGTCCCGGCGGAGGAGCAGAAATTGAAGTCGAAAAATACATT       555
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGAACATACCGGTGGTCCCGGCGGAGGAGCAGAAATTGAAGTCGAAAAATACATT       420

Qy   556 GTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGGCGCTGCTGCAGCTGAAA  615
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGGCGCTGCTGCAGCTGAAA  480

Qy   616 TCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACTGTGTGCCTTCCCCCG   675
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCGGTGGTCCGCACTGTGTGCCTTCCCCCG   540

Qy   676 GCGGACCTGCAGCTGCCGGACTGCCGGACTCTCCGGCTACGGCCAAGCATGAG         735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCGGACCTGCAGCTGCCGGACTGCCGGACTGTGAGCTCTCCGGCTACGGCCAAGCATGAG   600

Qy   736 GCCTTGTCTCCTTTCTATTCGGAGCGGCTCATGTCATGTGACCATCC              795
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCCTTGTCTCCTTTCTATTCGGAGCGGCTCATGTCATGTGACCATCC              660

Qy   796 AGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGACAACAGTGCTGTGTGCT     855
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AGCCGCTGCACATCACAACATTTAACAGAACAGTCACCGACAACAGTGCTGTGTGCT     720

Qy   856 GGAGACACTCGGAGCGGCGGC-GGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGCGATTC  914
         ||||||||||||||||||||| |||  |||||||||||||||||||||||||||||||||
Db   721 GGAGACACTCGGAGCGGCGGCGGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGCGATTC  780

Qy   915 GGGAGGCCCCCTGGTGTGTCTGAACGATGCCGCGAT-GACTTTGGTGGGCATCATCAGCT    973
         ||||  |||||||||||||||||||||||||||| | |||||||||||||||||||||||
Db   781 GGGAGGCCCCCTGGTGTGTCTGAACGATGCCGATGGCCATGCATCATCAGCT           840

Qy   974 GGGGCCTGGGCTGTGGACAGAAGGA                                      998
         ||||  ||||||||||||||||||
Db   841 GGGGGCCCTGGGCTGTGGGGGGACAGA                                    865

RESULT 5
BQ278177
LOCUS       BQ278177                 993 bp    mRNA    linear    EST 07-MAY-2002
DEFINITION  AGENCOURT_7061709 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5805153
```

```
ACCESSION   BQ278177
VERSION     BQ278177.1  GI:20488385
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 993)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/,
            National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished
JOURNAL     Contact: Robert Strausberg, Ph.D.
COMMENT     Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2043  row: o  column: 10
            High quality sequence stop: 681.
FEATURES    Location/Qualifiers
     source 1..993
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5805153"
            /tissue type="teratocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_109"
            /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
            XhoI; cDNA made by oligo-dT priming.  Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G).  Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            SuperScript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."
BASE COUNT  235 a  295 c  269 g  194 t
ORIGIN

Query Match             76.6%;  Score 815.8;  DB 13;  Length 993;
Best Local Similarity   99.6%;  Pred. No. 2.9e-202;
Matches   828;  Conservative  0;  Mismatches  2;  Indels  1;  Gaps  1;

Qy    234 GACGTGGGAGTACTGTGATGTGCCCTCCTCCTCCACCTGCCGGCTGAGACAGTACAGCCA 293
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      1 GACGTGGGAGTACTGTGATGTGCCCTCCTCCT-CACCTGCCGGCTGAGACAGTACAGCCA 59

Qy    294 GCCTCAGTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTCCCACCCCCTGGCAGGC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GCCTCAGTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTCCCACCCCCTGGCAGGC 119

Qy    354 TGGCCATCTTGCCAAGCACAGAGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACT 413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 TGGCCATCTTGCCAAGCACAGAGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACT 179

Qy    414 CATCAGCTCCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 CATCAGCTCCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 239

Qy    474 CCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCCCTGGCGGGAGGAGGAGCAGAA 533
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCCCTGGCGGGAGGAGGAGCAGAA 299

Qy    534 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGA 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAATGA 359

Qy    594 CATTGCGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCGTGGT 653
          ||||||||||||||||||||||||||||||||||||||||
Db    360 CATTGCGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCGTGGT 419

Qy    654 CCGCACTGTGTGCCTTCCCCGGGACCTGCAGCTGCCGGACCTGCAGCTCGGAGTGTGAGCT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CCGCACTGTGTGCCTTCCCCGGGACCTGCAGCTGCCGGACCTGCAGCTCGGAGTGTGAGCT 479

Qy    714 CTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCCTTCCTTCATTCGGAGCGGCTGAGGAGGC 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 CTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCCTTCCTTCATTCGGAGCGGCTGAGGAGGC 539

Qy    774 TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGT 833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGT 599

Qy    834 CACCGACAACATGTGTGTGCTGAGACACTCGGAGCGCGGGGCCCCAGGCAAACTTGCA 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CACCGACAACACATGTGTGTGCTGAGACACTCGGAGCGCGGGGCCCCAGGCAAACTTGCA 659

Qy    894 CGACGCCTGCCAGGGGATTCGGGAGGGCCCCGTGGTGTCTGAACGATGGCCGCCATGAC 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CGACGCCTGCCAGGGCGATTCGGGAGGGCCCCGTGGTGTCTGAACGATGGCCGCCATGAC 719

Qy    954 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGGACAAGAAGGATGTCCCGGGTGTGTA 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGGACAAGAAGGATGTCCCGGGTGTGTA 779

Qy   1014 CACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGGACCGTG 1064
          ||| |||||||||||||||||||||||||||||||||||||||||||||
Db    780 CACCAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGGACCGTG 830


RESULT 6
BQ690978
LOCUS       BQ690978  AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
            5', mRNA sequence.                883 bp    mRNA    linear    EST 15-JUL-2002
DEFINITION  AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
            5', mRNA sequence.
ACCESSION   BQ690978
VERSION     BQ690978.1  GI:21816294
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/,
            National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished
JOURNAL     Contact: Robert Strausberg, Ph.D.
COMMENT     Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2367  row: o  column: 09
            High quality sequence stop: 658.
FEATURES    Location/Qualifiers
     source 1..883
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6209120"
            /tissue type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_110"
            /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
            Site_2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
```

KEYWORDS   EST.
SOURCE   Homo sapiens (human)
             Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 918)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen. This sequence belongs to sequence cluster
             4498.r For more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_2&cluster=4498.r.
             Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_2.
FEATURES   Location/Qualifiers
source   1..918
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0DI039YK15"
             /tissue type="PLACENTA COT 25-NORMALIZED"
             /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
             /note="1st strand cDNA was primed with a NotI-oligo(dT)
             primer. Five prime end enriched, double-strand cDNA was
             digested with Not I and cloned into the Not I and EcoR V
             sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT   189 a   273 c   264 g   186 t   6 others
ORIGIN

Query Match   75.8%;   Score 807.6;   DB 13;   Length 918;
Best Local Similarity 96.4%; Pred. No. 4e-200;
Matches 855;   Conservative 0;   Mismatches 29;   Indels 3;   Gaps 3;

```
Qy 170 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 229
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  17 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  76

Qy 230 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTCCACCTGGCGCCCTGAGACAGTACA 289
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  77 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTCCACCTGGCGCCCTGAGACAGTACA 136

Qy 290 GCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 349
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 GCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 196

Qy 350 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCCCGGAGGAGCGGTTCCTGTGCGGGGGCA 409
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCCCGGAGGAGCGGTTCCTGTGCGGGGGCA 256

Qy 410 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCACTGCTTCCAGGAGGAGGTTTCCGC 469
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 257 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCACTGCTTCCAGGAGGAGGTTTCCGC 316

Qy 470 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGCCTCCACCCCTGGC 529
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 317 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGCCTCCACCCCTGGC 376

Qy 530 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 589
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 377 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 436

Qy 590 ATGACATTGCGCTGCTGCAGCTGAAATCGGAATTCGTCCCGTGTGCCCAGGAGCAGCAGCG 649
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 437 ATGACATTGCGCTGCTGCAGCTGAAATCGGAATTCGTCCCGTGTGCCCAGGAGCAGCAGCG 496

Qy 650 TGGTCCGCACTGTGTGCCTTCCCCCGGCGAGCTGGAGACAGTCGGACGGCGGACGGAGTGTG 709
```

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT   184 a   265 c   259 g   175 t   0;
ORIGIN

Query Match   76.2%;   Score 811.2;   DB 13;   Length 883;
Best Local Similarity 99.0%; Pred. No. 4.5e-201;
Matches 816;   Conservative 0;   Mismatches 8;   Indels 0;   Gaps 0;

```
Qy 112 GGCAAGGTTTACACAGCACCAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAAT 171
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGCAAGGTTTACACAGCACCAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATAAT  60

Qy 172 TACTGCCGGAATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 231
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 TACTGCCGGAATCCTGATGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy 232 CTGACGTGGGAGTACTGTGATGTGCCCTCGCTCCACCTGGCGCCCTGAGACAGTACAGC 291
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGACGTGGGAGTACTGTGATGTGCCCTCGCTCCACCTGGCGCCCTGAGACAGTACAGC 180

Qy 292 CAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAG 351
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAG 240

Qy 352 GCTGCCATCTTTGCCAAGCACAGGACCAGCAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATA 411
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GCTGCCATCTTTGCCAAGCACAGGACCAGCAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATA 300

Qy 412 CTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCC 471
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCC 360

Qy 472 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGTGGGCGAGCAG 531
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGTGGGCGAGCAG 420

Qy 532 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 591
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAT 480

Qy 592 GACATTGCGCTGCTGCAGCTGAAATCGGAATTCGTCCCGGATTCGTCCCGGAGCAGCAGGCTG 651
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GACATTGCGCTGCTGCAGCTGAAATCGGAATTCGTCCCGGATTCGTCCCGGAGCAGCAGGCTG 540

Qy 652 GTCCGCACTGTGTGCCTTCCCCCGGCGAGCTGCGGACTGGAGACGGACGGAGTGTGAG 711
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTCCGCACTGTGTGCCTTCCCCCGGCGAGCTGCGGACTGGAGACGGACGGAGTGTGAG 600

Qy 712 CTCTCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCGTGAAGGAG 771
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CTCTCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCGTGAAGGAG 660

Qy 772 GCTCATGTCAGACTGCCATCCAGCCGCTGCACATCACAACATTTACTTACTTAACAGAACA 831
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GCTCATGTCAGACTGTCATCAGACTGCCCATCACCCATCACAACATTTACTTACTTAACAGAACA 720

Qy 832 GTCACCGACAACATGCTGTGTCGTGGAGACACTCGGACGCGGCCCCCAGGCAAACTTG 891
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GTCACCGACAACATGCTGTGTCGTGGAGACACTCGGACGCGGCCCCCAGGCAAACTTG 780

Qy 892 CACGACGCGCTGCCAGGGGGCGATTCGGGGAGGCCCCCTGTGTGTCT 935
       |||||||||||||||||||||||||||||||||||||||||||||
Db 781 CACGACGCGCTGCCAGGGGGCGATTCGGGGAGGCCCCCTGTGTGTGT 824
```

RESULT 7
BX389608
LOCUS   BX389608   918 bp   mRNA   linear   EST 08-MAY-2003
DEFINITION BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389608
VERSION   BX389608.1 GI:30463469

ORIGIN

```
Query Match            74.8%;  Score 796.2;  DB 13;  Length 930;
Best Local Similarity  99.4%;  Pred. No. 3.8e-197;
Matches  820;  Conservative  0;  Mismatches  0;  Indels  2;  Gaps  2;

Qy    234 GAGGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGTCTGAGACAGTACAGCCA 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGTCTGAGACAGTACAGCCA 60

Qy    294 GCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGCAGGC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGCAGGC 120

Qy    354 TGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGGATACT 413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGGCATACT 180

Qy    414 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 240

Qy    474 CCACCTGACGGTGATCTCTGGGCAGAACATACCGGGTCCTGGCGAGGAGGAGCAGAA 533
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCACCTGACGGTGATCTTCTGGGCAGAACATACCGGGTCCTGGCGAGGAGGAGCAGAA 300

Qy    534 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 593
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 360

Qy    594 CATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGT 653
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGT 420

Qy    654 CCGCACTGTGTGCCTTCCCCCGCGGACCTGACCTGGACGTGGACGGAGTGTGAGCT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCGCACTGTGTGCCTTCCCCCGCGGACCTGACCTGGACGTGGACGGAGTGTGAGCT 480

Qy    714 CTCGGCTACGGCAAGCATGAGGCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGC 773
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCGGCTACGGCAAGCATGAGGCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGC 540

Qy    774 TCATGTCAGACTGTACCCATCCAGCGCGTCGACATCACAACATTTACTTAACAGAACAGT 833
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCATGTCAGACTGTACCCATCCAGCGCGTCGACATCACAACATTTACTTAACAGAACAGT 600

Qy    834 CACCGACAACATGCTGTGTGCTGAGACACTCGGAGCCGCGGCCCCAGGCAAACTTGCA 893
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CACCGACAACATGCTGTGTGCTGAGACACTCGGAGCCGCGGCCCCAGGCAAACTTGCA 660

Qy    894 CGACGCCTGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTGTGTCTGAACGATGGCCGCATGAC 953
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGACGCCTGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTGTCTGAACGATGCCCGCATGAC 720

Qy    954 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGGACAGAAGGATGTCCC-GGGTGTGT 1012
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGGACAGAAGGATGTCCCGGGGTGTGT 780

Qy   1013 ACA-CAAAGGTTACCAACTACCTTAGACTGGATTCGTGACAACATG 1056
          ||| ||||||||||||||||||||||||||||||||||||||||||
Db    781 ACACCCAAGGGTACCAACTACCTTAGACTGGATTCGTGACAACATG 825
```

RESULT 9
BU157720

LOCUS   BU157720   888 bp   mRNA   linear   EST 04-SEP-2002
DEFINITION AGENCOURT_8043192 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084476
            5', mRNA sequence.
ACCESSION   BU157720
VERSION   BU157720.1   GI:22671252
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens

```
Db    497 TGGTCCGACACTGTGTGCCTTCCCCCGCGGGACCTGACCTGGACGGAGTGTG 556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    710 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTCTATTCGGAGCGGCGCTGAAGG 769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 616

Qy    770 AGGCTCATGTCAGACTGTACCCATCCAGCGCGTCGACATCACAACATTTACTTAACAGAA 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    617 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGTCGACATCACAACATTTACTTAACANGA 676

Qy    830 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 889
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    677 CAGTCACCGACANCACATGCTGTGTGCTGAGACACTCGNAGCGGCGGCGGCCCCA-GCAAACT 735

Qy    890 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 949
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 TGCACGACGCCTGCCAGGGCGATTNCGGAGGCCCCCTGGTGTGTCTGAACGATTGCCCAA 795

Qy    950 TGA-CTTTGGTGGGCATCATCAGCTGGGG--CCTGGGCGTGGACAGAAGGATGTCCCGGG 1007
          ||| ||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db    796 TGACCTTTGGTGGTGCCATCATCAGCTGGGG   CCTGGGCGTGTTGACAGAAGGATGTTCCCGG 855

Qy   1008 TGTGTACACAAAGGTTACCAACTACCTTAGACTGGATTCGTGACAAACA 1054
          |||| | ||| ||||||||||||||||||||||||||||||||||||||
Db    856 TGTGTCACCCAAGTTTACAACTACCTTACCTTAGACTGGATTCGGGGACAACA 902
```

RESULT 8
BU845014

LOCUS   BU845014   930 bp   mRNA   linear   EST 16-OCT-2002
DEFINITION AGENCOURT_10412263 NIH_MGC_109 Homo sapiens cDNA clone
            IMAGE:6578533 5', mRNA sequence.
ACCESSION   BU845014
VERSION   BU845014.1   GI:24029455
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2780 row: g column: 13
            High quality sequence stop: 639.
FEATURES   Location/Qualifiers
source   1..930
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6578533"
                /tissue_type="teratocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_109"
                /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
                XhoI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCACGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH_MGC Library."
BASE COUNT   199 a   279 c   266 g   186 t

```
Qy   592  GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCGTGTGCCAGGAGAGCAGCGTG 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCGTGTGCCAGGAGAGCAGCGTG 540

Qy   652  GTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGCCGGACTGCCGAGTGTGAG 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGCCGGACTGCCGAGTGTGAG 600

Qy   712  CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAG 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAG 660

Qy   772  GCTCATGTCAGACTGTACCCATCCAGCCCGCACATCACACATTTACTTAACAGAACA 831
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GCTCATGTCAGACTGTACCCATCCAGCCCGCACATCACACATTTACTTAACAGAACA 720

Qy   832  GTCACCGACAAACATGCTGT-GTGCTGGAGACACTCGG--AGCGCGGGCCCCAGGCAAAC 888
          |||||||||||||||||||| |||||||||||||||||  |||||||| ||||||||||||
Db   721  GTCACCGACAACATGCTGTGGTGCTGGAGACACTCGGGAGGACGCGGGGCCCCAGCCAAAC 780

Qy   889  TTGCACGACGCCT-GCCAGGGGGCGATTCGGGAGGCCCCCGGT-GTGTCGAACGATGG-C 945
          ||||||||||||| ||| |||||||||||||||||||||||||| ||||||||||||||| |
Db   781  TTGCACGACGCCTGGCCGGGGCGATTCGGGAGGCCCCCGGTGGTCGGTGGGGTCTGAACCATGGCC 840

Qy   946  CGCATGACTTT--GGTGGGCATCATCAGCTGGGGCCTGGGCTGTGG 989
          ||||||||||   ||||| |||||||||||||||| |||||||||
Db   841  CGCATGACTTTTGGGGGGCATCATCAACTGGGGCGCTGGGGCTGG 886
```

```
RESULT 10
BI765247
LOCUS       BI765247                  916 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603050338F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190656 5',
            mRNA sequence.
ACCESSION   BI765247
VERSION     BI765247.1  GI:15756825
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 916)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1476  row: k  column: 09
            High quality sequence stop: 837.
FEATURES             Location/Qualifiers
     source          1..916
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5190656"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female. Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning). Average insert size 1.4 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was constructed by C.
```

Left column:

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2316  row: e  column: 21
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1..888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6084476"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT  181 a    265 c    270 g    172 t
ORIGIN
```

```
Query Match           73.5%;  Score 782.8;  DB 13;  Length 888;
Best Local Similarity 97.2%;  Pred. No. 1.2e-193;
Matches 861;  Conservative  0;  Mismatches  17;  Indels  8;  Gaps  6;

Qy   112  GGCAAGGTTACACAGCACAGAACCCCAGTGCCCAGGCCACTCGGCCTGGCAAACATAAT 171
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GGCAAGGTTACACAGCACAGAACCCCAGTGCCCAGGCCACTGGGCCTGGCAAACATAAT 60

Qy   172  TACTGCCCGGAATCCGATGGGGATGCCAAGCCCCTGGTGCCACGTGCTGAAGAACCGCAGG 231
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TACTGCCCGGAATCCGATGGGGATGCCAAGCCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy   232  CTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGGCGGCCTGAGACAGTACAGC 291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGGCGGCCTGAGACAGTACAGC 180

Qy   292  CAGGCCTCAGTTCGCATCAAAGGAGGGGCTCTTCGCCGACGATCGCTCCCACCCCTGCCAG 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CAGCCCTCAGTTCGCATCAAAGGAGGGGCTCTTCGCCGACGATCGCTCCCACCCCTGCCAG 240

Qy   352  GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCCGGAGACGGGTTCCTGTGCGGGGCATA 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCCGGAGACGGGTTCCTGTGCGGGGGCATA 300

Qy   412  CTCATCAGCTCCTGCTGGATTCTCTGCGCGCCCACTGTTCCAGGAGAGAGGTTTCCGCCC 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CTCATCAGCTCCTGCTGGATTCTCTGCGCGCCCACTGTTCCAGGAGAGAGGTTTCCGCCC 360

Qy   472  CACCACCTGACCGGTGACTTGGGCAGAACATACCGGGTGGTCCTGGCAGGAGGAGGGCAG 531
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CACCACCTGACCGGTGACTTGGGCAGAACATACCGGGTGGTCCTGGCAGGAGGAGGGCAG 420

Qy   532  AAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTACGACAAT 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTACGACAAT 480
```

```
                    Gruber (Invitrogen). Research Genetics tracking code
                    023. Note: this is a NIH_MGC Library."
BASE COUNT    204 a    272 c    256 g    184 t
ORIGIN

Query Match    72.1%;    Score 767.4;  DB 12;  Length 916;
Best Local Similarity 99.7%;    Pred. No. 1.3e-189;
Matches  779;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  1;

Qy  286  TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCC  345
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCC   60

Qy  346  TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCTGGAGAGCGGTTCCTGTGCCGG  405
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCTGGAGAGCGGTTCCTGTGCCGG  120

Qy  406  GGCATACTCATCAGCTCCTGCTGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTT  465
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GGCATACTCATCAGCTCCTGCTGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTT  180

Qy  466  CCGCCCACCACCTGACGGTGATCTTGGGCAGAACATACGGGTGGTCCCTGGCGAGGAG  525
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACGGGTGGTCCCTGGCGAGGAG  240

Qy  526  GAGCAGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTAC  585
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  GAGCAGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTAC  300

Qy  586  GACAATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGAGC  645
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GACAATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGAGC  360

Qy  646  AGCGTGGTCCGGACACTGTGTGCCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAG  705
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  AGCGTGGTCCGGACACTGTGTGCCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAG  420

Qy  706  TGTGAGCTCTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTG  765
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  TGTGAGCTCTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTG  480

Qy  766  AAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACAATCACACAGATTTACTTAAC  825
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  AAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACAATCACACAGATTTACTTAAC  540

Qy  826  AGAACAGTCACCGACAACATGCTGTGTCTGAGACACTCGGAGGCGCGGGCCCCAGGCA  885
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  AGAACAGTCACCGACAACATGCTGTGTCTGAGACACTCGGAGGCGCGGGCCCCAGGCA  600

Qy  886  AACTTGCACGACGCCCTGCCCAGGGGCGATTCGGAGGGCCCCCTGGTGTCTGAACGATGGC  945
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  AACTTGCACGACGCCCTGCCCAGGGGCGATTCGGAGGGCCCCCTGGTGTCTGAACGATGGC  660

Qy  946  CGCATGACTTTGTGGGGCATCATCAGCTGGGGCGTGTGGACAGAAGGATGTCCCG  1005
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  CGCATGACTTTGTGGGGCATCATCAGCTGGGGCGTGTGGACAGAAGGATGTCCCG   720

Qy 1006  GGTGTGTACACAAAGGTTACCAACTACCTAGAC-TGGATTGCTGACAACATGCGACCGTG  1064
         |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  721  GGTGTGTACACCAAGGTTACCAACTACCTAGACTTGGATTGCTGACAACATGCGACCGTG   780

Qy 1065  A  1065
         |
Db  781  A   781
```

RESULT 11
BQ927899
LOCUS   BQ927899   950 bp   mRNA   linear   EST 20-AUG-2002
DEFINITION AGENCOURT_8775988 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6373489
            5', mRNA sequence.
ACCESSION   BQ927899
VERSION   BQ927899.1 GI:22342930

FEATURES   Location/Qualifiers
     source   1..950
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6373489"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT   207 a   283 c   279 g   181 t
ORIGIN

Query Match   71.8%;   Score 765.2; DB 13; Length 950;
Best Local Similarity 98.1%;   Pred. No. 4.9e-189;
Matches 806; Conservative 0; Mismatches 13; Indels 13; Gaps 3;

```
Qy  210  CCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCAC  269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43  CCCCGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCAC  102

Qy  270  CTGCGGCCTGAGACAGTACAGCCAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGGA  329
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  103  CTGCGGCCTGAGACAGTACAGCCAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGGA  162

Qy  330  CATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGA  389
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  163  CATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGA  222

Qy  390  GCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCTGCTGGATTCTCTGCTCCGCCCACTG  449
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  223  GCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCTGCTGGATTCTCTGCTCCGCCCACTG  282

Qy  450  CTTCCAGGAGAGG-TTTCCGCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGG  508
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  283  CTTCCAGGAGAGGTTTTCCGCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGG  342

Qy  509  TGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTGTCCATAAGGAAT  568
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  343  TGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAAT  402

Qy  569  TCGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCC  628
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  403  TCGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCC  462
```

```
Qy   629  GCTGTGCCCAGGAGCAGCGCGTGTCCGCACTGTGTGCCTTCCCCGGCGGCGGACCTGCAGC  688
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   463  GCTGTGCCCAGGAGCAGCGCGTGTCCGCACTGTGTGCCTTCCCCGGCGGCGGACCTGCAGC  522

Qy   689  TGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTT   748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   523  TGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCCTT   582

Qy   749  TCTATTCGGAGCGGCTGAAGGAGGAGGCTCATGTCGAGACTGTACCATCCAGCCGCTGCACAT  808
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   583  TCTATTCGGAGCGGCTGAAGGAGGAGGCTCATGTCGAGACTGTACCATCCAGCCGCTGCACAT  642

Qy   809  CACAACATTTACTTAACAGAACACAGTCACCGACACACATGCTGTGTGCTGGAGACACTCCGA  868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   643  CACAACATTTACTTAACAGAACACAGTCACCGACACACATGCTGTGTGCTGGAGACACTCGGA  702

Qy   869  GCGGCGGGCCCCAGGCAAACTTGCACGACGCGCCTGCCAGGGCGATTCGGGAGGCCCCTGG  928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   703  GCGGCGGGCCCCAGGCAAACTTGCACGACGCGCCTGCCAGGGCGATTCGGGAGGCCCCTGG  762

Qy   929  TGTGTCTGAACGATGGCCGCATGACTTTGGTGGTGGCATCATCAG-CTGGGGCCTGGGCTGT  987
          |||||||||||||||||||||||||||||||||||||| ||| |||||||||||||
Db   763  TGTGTCTGAACGATGGCCGCATGACTTTGGTGGTGGCATCATCAGCCTGGGGCCTGGGCTGT  822

Qy   988  GGACAG-AAGGATGTCCCGGGTGTGTACACAAAGGTTACCAA  1028
          ||||| ||| |||||| ||||||||| ||| ||||
Db   823  GGACAGAAAGAAGAAGCCCCGGGTGGGGACCACCAAGGGTACCCA  864
```

```
BASE COUNT      173 a      295 c      293 g      187 t
ORIGIN
```

```
Query Match           66.4%;  Score 706.8;  DB 10;  Length 948;
Best Local Similarity 93.3%;  Pred. No. 8.9e-174;
Matches  793;  Conservative   0;  Mismatches   52;  Indels   5;  Gaps   5;
```

```
Qy     6  GGGAAACAGTGACTGCTACT-TTGGGAATAGGGTCAGCCTACCGTGGCACGCACAGCCTCA   64
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     8  GAGGAACAGTGACTGCTATTATTGGGAATAGGGTCAGCCTACCGTGGCACGCACAGCCTCA   67

Qy    65  CCGAGTCGGTGCCTCCTGCCTCTGCCTCCC-GTGGAATTCCATGATCCTGATAGGCAAGGTTTAC  123
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    68  CCGAGTCGGTGCCTCCTGCCTCTGCCTCCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTAC  127

Qy   124  ACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAAT  183
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128  ACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAAT  187

Qy   184  CCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCTGAAGAACCGCAGGCTGACTGGGAG  243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188  CCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCTGAAGAACCGCAGGCTGACTGGGAG  247

Qy   244  TACTGTGATGTGCCCTCCTGCTCCACCTGTCGGGCCTGAGACAGTACAGCCAGCCTCAGTT  303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248  TACTGTGATGTGCCCTCCTGCTCCACCTGTCGGGCCTGAGACAGTACAGCCAGCCTCAGTT  307

Qy   304  CGGATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGCCTGCCATCTTT  363
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308  CGGATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGCCTGCCATCTTT  367

Qy   364  GCCAAGCACACAGGAGTGGGGTCCCGGAGAGGCGGGTTCCTGTGGGGGGGGCATACTCATCAGCTCC  423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368  GCCAAGCACACAGGAGTGGGGTCCCGGAGAGGCGGGTTCCTGTGGGGGGGGCATACTCATCAGCTCC  427

Qy   424  TGCTGGATTCTCTGCGCGCCACTGCTTCCAGGAGGAGAGTTCGCCCCCACCACCTGACG  483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428  TGCTGGATTCTCTGCGCGCCACTGCTTCCAGGAGGAGAGTTCGCCCCCACCACCTGACG  487

Qy   484  GTGATCTTGGGCAGAACATACCGGGTGGTCCCGGTGGTCCCTGGCGAGGAGGCAGAAATTTGAAGTC  543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488  GTGATCTTGGGCAGAACATACCGGGTGGTCCCGGTGGTCCCTGGCGAGGAGGCAGAAATTTGAAGTC  547

Qy   544  GAAAAATACATTGTCCATAAGGAATTCGATGATGATGAACACTTACGACAATGACAATTGCGCTG  603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548  GAAAAATACATTGTCCATAAGGAATTCGATGATGATGAACACTTACGACAATGACAATTGCGCTG  607

Qy   604  CTGCAGCTGAAATCGGATTCGTCCGACTGGCTCGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTG  663
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608  CTGCAGCTGAAATCGGATTCGTCCGACTGGCTCGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTG  667

Qy   664  TGCCTTCCCCGGCGGACGACCTGCAGCTGCGCAGCTGACCGGAGTGTGAGCTCTCCGGCTAC  723
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668  TGCCTT-CCCCGGCGGACGACCTGCAGCTGCGCAGCTGACCGGAGTGTGAGCTCTCCGGCTAC  726

Qy   724  GGCAAGCATGAGGCCTTGTCTCCTTTCTCATTCGGAGCGGCTGAAGGAGGAGGCTCATGTCAGA  783
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   727  GGCAAGCATGAGG--CTTGTCTCTTTTTCATTCGGAGCCGGCTGAAGGGGCTCATGTCGGA  785

Qy   784  CTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAAC  843
          |||||||||||||||| |||| ||| |||||||||||||||||||||||||||||||||||
Db   786  CTG-GTCCATCCAGCCGCTGGGCTTCCCCACCTTTCTTCAGGACAGCAGCAGCCGCCTGC  844

Qy   844  ATGCTGTGTG  853
          ||||||||||
Db   845  TGGGGCTGGG  854
```

ACCESSION   BG763582
VERSION   BG763582.1 GI:14074235
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 958)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1719 row: n column: 21
             High quality sequence stop: 810.
FEATURES   Location/Qualifiers
     source   1..958
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4861268"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_49"
                     /note="Organ: sKin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT   198 a   285 c   282 g   193 t
ORIGIN

Query Match 65.5%;   Score 697.4;   DB 12;   Length 958;
Best Local Similarity 94.2%;   Pred. No. 2.6e-171;
Matches 834;   Conservative 0;   Mismatches 36;   Indels 15;   Gaps 10;

Qy   1 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   72 TCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 131

Qy   61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGTT 191

Qy   121 TACACAGCACAGAACCCCAGTGCCAGGACTGGGCCTGGCAAACATAATATTACTGCCGG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 TACACAGCACAGAACCCCAGTGCCAGGACTGGGCCTGGCAAACATAATATTACTGCCGG 251

Qy   181 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 AATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 311

Qy   241 GAGTACTGTGATGTGCCCTTCGTCCTGCCGGCCTCCACCTGGGCCGTACAGCCAGCCTCAG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   312 GAGTACTGTGATGTGCCCTTCGTCCTGCCGGCCTCCACCTGGGCCGTACAGCCAGCCTCAG 371

Qy   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTTCCACCCCTGCCAGGCTGCCATC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTTCCACCCCTGCCAGGCTGCCATC 431

Qy   361 TTTGCCAAGCAGCACAGGAGGTCGCCCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCATCAGC 420

Db   432 TTTGCCAAGCACAGGAGGTCGCCCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCATCAGC 491

Qy   421 TCCTGCTGGATTCTCTGCTGCCGCCCCACTGCTTCCAGGAGGAGTTCCGCCCACCACCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   492 TCCTGCTGGATTCTCTGCTGCCGCCCCACTGCTTCCAGGAGGAGTTCCTCTGCCCACCACCTG 551

Qy   481 ACGGTGATCTTGGGCAGAACATACCCGGTGGTCCCTGGCGAGGAGCAGAGAAATTTGAA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 ACGGTGATCTTGGGCAGAACATACCCGGTGGTCCCTGGCGAGGAGCAGAGAAATTTGAA 611

Qy   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   612 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACAACTTACGACAATGACACTTGCG 671

Qy   601 CTGCTGCAGCTGAAATCGGATTCGT-CCCGCTGTGCCCAGGAGCAGC-GTGGTCCGCA 658
         ||||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||
Db   672 CTGCTGCAGCTGAAATCGGATTCGTCCCCGCTGTGCCCAGGAGCAGCAGCGGTGGTCCGCA 731

Qy   659 CGTGTGCCTT--CCCCCGGCGGACCTGCAGCTGCCGCTGCCGACGACTGAGCTCT 715
         ||||||||| ||||||||||||||||||||||||||||||| |||||||||||||
Db   732 CGTGTGCCTTTCCCCCGGCGGACCTGCAGCTGCCGCTGCCGGACTGAGCTGTGAGCTCT 791

Qy   716 CCG-GCTACGGCAAGCAT--GAGGCCTTGTCTGCTCCTTTCTATTCCGAGCGGCTGAAGGAGG 772
         ||| ||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   792 CCGAGCTACGGCAAGCATTGAGGGCCTTGTCTGCTCCTTTCTATTCCGAGCGGCCGG-TGAAGGAGG 850

Qy   773 CTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAG 832
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   851 CTCATG--CGACTGGACCCATTCCACCGGTGGCCATTCCAACCTTTACTT-ACAGAACAG 907

Qy   833 TCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCCGCGGGC 877
         ||||||||||||||||||||||||| |||||||||||||||||||
Db   908 TACC--GAAAATGTGTGTGCGGGGACCCCCGACGGGGCCCAGGC 949

RESULT 14
BQ960216
LOCUS   BQ960216
DEFINITION AGENCOURT_8945604 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6459629
            5', mRNA sequence.
ACCESSION   BQ960216
VERSION   BQ960216.1 GI:22375694
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 974)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2645 row: a column: 06
            High quality sequence stop: 603.
FEATURES   Location/Qualifiers
     source   1..974
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6459629"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_101"
                     /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:

VERSION   BE732704.1 GI:10146696
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 709)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
            National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE   Unpublished
  JOURNAL
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM752 row: d column: 17
            High quality sequence start: 3
            High quality sequence stop: 701.

FEATURES   Location/Qualifiers
     source   1..709
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3925600"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_21"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). size-selected >500bp
                     for average insert size 1.8kb. Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT   147 a   219 c   203 g   140 t
ORIGIN

Query Match   65.1%; Score 693.4; DB 10; Length 709;
Best Local Similarity 99.7%; Pred. No. 2.5e-170;
Matches 705; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

Qy   197 CCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGC 256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGC 60

Qy   257 CCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAG 316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAG 120

Qy   317 GGCTCTTCGCCGACATCGCCTCCACCCTGCCAGGTGCCATCTTTGCCAAGCACAGGA 376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGCTCTTCGCCGACATCGCCTCCACCCTGCCAGGTGCCATCTTTGCCAAGCACAGGA 180

Qy   377 GGTGCGCCCGGAGAGCCGGTTCCTGTGTGCTCCGGGGCATGAGCCATCATCATCAGCTCT 436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGTGCGCCCGGAGAGCCGGTTCCTGTGTGCTCCGGGGCATGAGCCATCATCATCAGCTCT 240

Qy   437 CTGCCGGCCCACTGCTTCCAGGAGCAAGCATGGAGGCCTTGTTCCTGTCGCCGCCGGA 496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGCCGGCCCACTGCTTCCAGGAGCAAGCATGGAGGCCTTGTTCCTGTCGCCGCCGGA 300

Qy   497 GAACATACCGGGTGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTG 556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAACATACCGGGTGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTG 360

Qy   557 TCCATAAGGAATTCGATGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAAT 616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCATAAGGAATTCGATGATGATGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAAT 420

            XhoI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Library constructed by Ling Hong (University of California,
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."

BASE COUNT   215 a   293 c   279 g   187 t
ORIGIN

Query Match   65.2%; Score 694.4; DB 13; Length 974;
Best Local Similarity 97.8%; Pred. No. 1.6e-170;
Matches 757; Conservative   0; Mismatches 11; Indels   6; Gaps   5;

Qy   276 CCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGC 335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 CCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGC 78

Qy   336 CTCCACCCCTGGCCAGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTT 395
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   79 CTCCACCCCTGGCCAGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTT 138

Qy   396 CCTGTGCGGGGGCATACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCA 455
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   139 CCTGTGCGGGGGCATACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCC-CTGCTTCCA 197

Qy   456 GGAGAGGTTTCCGCCCCACCACTGACGTGATCTTGGGCAGAACATACCGGGTGGTCCC 515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 GGAGAGGTTTCCGCCCCACCACTGACGTGATCTTGGGCAGAACATACCGGGTGGTCCC 257

Qy   516 TGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTGTCCATAAGGAATTCGATGA 575
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 TGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTGTCCATAAGGAATTCGATGA 317

Qy   576 TGACACTTACGACAATGACATTGCGCTGCTGCGAGCTGAAATGGATTCGTCCCGCTGTGC 635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 TGACACTTACGACAATGACATTGCGCTGCTGCGAGCTGAAATGGATTCGTCCCGCTGTGC 377

Qy   636 CCAGGAGGAGCAGCGCTGGTCCGGCACTGTGTGCCTTCCCCGGCGACCTGCCGGA 695
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   378 CCAGGAGGAGCAGCGCTGGTCCGGCACTGTGTGCCTTCCCCGGCGACCTGCCGGA 437

Qy   696 CTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCAAGCATGGAGGCCTTGTCTTCTTATTC 755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCAAGCATGGAGGCCTTGTCTTCTTATTC 497

Qy   756 GGAGGCGGCTGAAGGAGGAGAGGCTCATGTCAGACTGTACCGATCGACACTCACACAACA 815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498 GGAGGCGGCTGAAGGAGGAGAGGCTCATGTCAGACTGTACCGATCGACACTCACACAACA 557

Qy   816 TTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGTCGGAGACACTCGGAGGCGCGG 875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 TTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGTCGGAGACACTCGGAGGCGCGG 617

Qy   876 GCCCCAGGCAAACTTGACGACGCTGCCAGACGGCTGCGACGGGAGG-CCCCGTGGTGTC 934
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   618 GCCCCAGGCAAACTTGACGACGCTGCCAGACGGCTGCGACGGGAGGGCCCCCGTGGTGTC 677

Qy   935 TGAACGATGGCCGCAT-GACTTTGGTGGGTACACAAGGTTACCAACTACTTAGACTGGATT 992
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |
Db   678 TGAACGATGGCCGCATGGACTTTGGTGGGTACACAAGGTTACCAACTACTTAGACTGGAAT 737

Qy   993 GAAGGATGTCCC--GGGTGTGTGTACACAAGGTTACCAACTACTTAGACTGGATT 1044
          |||||||| ||| |||||| ||||||||||||||||||||||||||| ||||||
Db   738 GAAGGATGGCCCGGGGGGGGGTACCCGGGGGTACCAACTACCTAAACTGAAT 791

RESULT 15
BE732704
LOCUS   BE732704   709 bp   mRNA   linear   EST 15-SEP-2000
DEFINITION 601571366F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925600 5',
            mRNA sequence.
ACCESSION   BE732704

Wed Aug 20 09:29:41 2003

```
Qy    617 CGGATTGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCGG   676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCGG   480

Qy    677 CGGACCTGCAGCTGCCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGG   736
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGGACCTGCAGCTGCCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGG   540

Qy    737 CCTTGTCTCCTTCTATTCGGAGCGGCTGAAGAGGGCTCATGTCAGACTGTACCATCCA   796
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTTGTCCCCTTTCTATTCGGAGCGGCTGAAGAGGGCTCATGTCAGACTGTACCATCCA   600

Qy    797 GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG   856
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    601 GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG   660

Qy    857 GAGACACTCGGAGCGGCGGGGGGCCCCAGGCAAACTTGCACGACGCCTGC   903
          ||||||||||||||||| |||||||| ||||||||||||||||||||||
Db    661 GAGACACTCGGAGCGCGGCGGGG-CCCAGGCAAACTTGCACGACGCCTGC   706
```

Search completed: August 19, 2003, 08:29:17
Job time : 3116.64 secs

THIS PAGE BLANK (usp10)

OM nucleic - nucleic search, using sw model

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 19:53:48 ; Search time 75.5794 Seconds
(without alignments)
9743.047 Million cell updates/sec

Title: US-09-987-455-1
Perfect score: 18
Sequence: 1 tctgagggaaacagtgac 18
Scoring table: IDENTITY.NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	100.0	18	6	AX430833	AX430833 Sequence
2	18	100.0	18	6	AX431267	AX431267 Sequence
3	18	100.0	42	6	AX430834	AX430834 Sequence
4	18	100.0	42	6	AX431288	AX431288 Sequence
5	18	100.0	952	6	I09039	I09039 Sequence 2
6	18	100.0	1065	6	A27447	A27447 DNA sequenc
7	18	100.0	1065	6	AR059995	AR059995 Sequence
8	18	100.0	1065	6	AX431270	AX431270 Sequence
9	18	100.0	1065	6	AX431273	AX431273 Sequence
10	18	100.0	1065	6	E01943	E01943 Synthetic D
11	18	100.0	1068	6	A27431	A27431 DNA sequenc
12	18	100.0	1068	6	A27433	A27433 DNA sequenc
13	18	100.0	1068	6	A27439	A27439 DNA sequenc
14	18	100.0	1068	6	A27445	A27445 DNA sequenc
15	18	100.0	1068	6	AR059987	AR059987 Sequence
16	18	100.0	1068	6	AR059988	AR059988 Sequence
17	18	100.0	1068	6	AR059991	AR059991 Sequence
18	18	100.0	1068	6	AR059994	AR059994 Sequence
19	18	100.0	1068	6	E01935	E01935 Synthetic D
20	18	100.0	1068	6	E01936	E01936 Synthetic D
21	18	100.0	1068	6	E01939	E01939 Synthetic D
22	18	100.0	1068	6	E01942	E01942 Synthetic D
23	18	100.0	1068	6	E05897	E05897 DNA sequenc
24	18	100.0	1068	6	E05898	E05898 DNA sequenc
25	18	100.0	1128	6	AX431268	AX431268 Sequence
26	18	100.0	1128	6	AX431271	AX431271 Sequence
27	18	100.0	1170	6	A27727	A27727 DNA sequenc
28	18	100.0	1170	6	A27729	A27729 DNA sequenc
29	18	100.0	1170	6	AR059997	AR059997 Sequence
30	18	100.0	1170	6	AR059998	AR059998 Sequence
31	18	100.0	1170	6	E01945	E01945 Synthetic D
32	18	100.0	1170	6	E01946	E01946 Synthetic D
33	18	100.0	1185	9	AF260825	AF260825 Homo sapi
34	18	100.0	1314	6	A27435	A27435 DNA sequenc
35	18	100.0	1314	6	A27437	A27437 DNA sequenc
36	18	100.0	1314	6	A27441	A27441 DNA sequenc
37	18	100.0	1314	6	AR059989	AR059989 Sequence
38	18	100.0	1314	6	AR059990	AR059990 Sequence
39	18	100.0	1314	6	AR059992	AR059992 Sequence
40	18	100.0	1314	6	AR059993	AR059993 Sequence
41	18	100.0	1314	6	E01937	E01937 Synthetic D
42	18	100.0	1314	6	E01938	E01938 Synthetic D
43	18	100.0	1314	6	E01940	E01940 Synthetic D
44	18	100.0	1314	6	E01941	E01941 Synthetic D
45	18	100.0	1341	6	A30593	A30593 DNA for tis

ALIGNMENTS

RESULT 1
AX430833
LOCUS   AX430833   18 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 5 from Patent WO0240696.
ACCESSION   AX430833
VERSION   AX430833.1   GI:21655910
KEYWORDS
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tavapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale protein production in prokaryotes
  JOURNAL   Patent: WO 0240696-A 5 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)

Pred. No. is the number of results predicted by chance to have a

```
FEATURES
     source          Location/Qualifiers
                     1..18
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence of the N-terminal part of K2S
                     molecule"
BASE COUNT        6 a        3 c        6 g        3 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 6;   Length 18;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db        1 TCTGAGGGAAACAGTGAC 18


RESULT 2
AX431267
LOCUS           AX431267          18 bp    DNA         linear   PAT 28-JUN-2002
DEFINITION      Sequence 1 from Patent WO0240650.
ACCESSION       AX431267
VERSION         AX431267.1  GI:21656149
KEYWORDS        .
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1
  AUTHORS       Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE         Methods for large scale production of recombinant dna-derived tpa
                or k2s molecules
  JOURNAL       Patent: WO 0240650-A 1 23-MAY-2002;
                BOEHRINGER INGELHEIM INT (DE)
FEATURES        Location/Qualifiers
     source     1..18
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
                /note="coding sequence of N-terminal part of K2S protein"
BASE COUNT        6 a        3 c        6 g        3 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 6;   Length 18;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db        1 TCTGAGGGAAACAGTGAC 18


RESULT 3
AX430834
LOCUS           AX430834          42 bp    DNA         linear   PAT 28-JUN-2002
DEFINITION      Sequence 6 from Patent WO0240696.
ACCESSION       AX430834
VERSION         AX430834.1  GI:21655911
KEYWORDS        .
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1
  AUTHORS       Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE         Methods for large scale protein production in prokaryotes
  JOURNAL       Patent: WO 0240696-A 6 23-MAY-2002;
                BOEHRINGER INGELHEIM INT (DE)
FEATURES        Location/Qualifiers
     source     1..42
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
```

```
BASE COUNT       10 a        9 c       19 g        4 t
ORIGIN                    /note="oligonucleotid sequence"

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 6;   Length 42;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db       25 TCTGAGGGAAACAGTGAC 42


RESULT 4
AX431288
LOCUS           AX431288          42 bp    DNA         linear   PAT 28-JUN-2002
DEFINITION      Sequence 22 from Patent WO0240650.
ACCESSION       AX431288
VERSION         AX431288.1  GI:21656157
KEYWORDS        .
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1
  AUTHORS       Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE         Methods for large scale production of recombinant dna-derived tpa
                or k2s molecules
  JOURNAL       Patent: WO 0240650-A 22 23-MAY-2002;
                BOEHRINGER INGELHEIM INT (DE)
FEATURES        Location/Qualifiers
     source     1..42
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
                /note="primer"
BASE COUNT       10 a        9 c       19 g        4 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 6;   Length 42;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db       25 TCTGAGGGAAACAGTGAC 42


RESULT 5
I09039
LOCUS           I09039           952 bp    DNA                  PAT 02-DEC-1994
DEFINITION      Sequence 2 from Patent WO 8808878.
ACCESSION       I09039
VERSION         I09039.1  GI:588252
KEYWORDS        .
SOURCE          Unknown.
  ORGANISM      Unknown.
                Unclassified.
REFERENCE       1  (bases 1 to 952)
  AUTHORS       Reddy,V.B.
  TITLE         SUBSTANCES FOR DETECTING BLOODCLOTS
  JOURNAL       Patent: WO 8808878-A 2 17-NOV-1988;
FEATURES        Location/Qualifiers
     source     1..952
                /organism="unknown"
BASE COUNT      225 a      266 c      286 g      175 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 6;   Length 952;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
```

Db   638 TCTGAGGGAAACAGTGAC 655

RESULT 6
A27447
```
LOCUS       A27447                  1065 bp    DNA     linear   PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pthTTrp.
ACCESSION   A27447
VERSION     A27447.1  GI:21727235
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1065)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 47 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1065
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      225 a    314 c    311 g    215 t
ORIGIN
```

Query Match   100.0%; Score 18; DB 6; Length 1065;
  Best Local Similarity   100.0%; Pred. No. 1.1e+02;
  Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db          4 TCTGAGGGAAACAGTGAC 21
```

RESULT 7
AR059995
```
LOCUS       AR059995                1065 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 60 from patent US 5840533.
ACCESSION   AR059995
VERSION     AR059995.1  GI:5986445
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1065)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 60 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1..1065
                     /organism="unknown"
BASE COUNT      225 a    314 c    311 g    215 t
ORIGIN
```

Query Match   100.0%; Score 18; DB 6; Length 1065;
  Best Local Similarity   100.0%; Pred. No. 1.1e+02;
  Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db          4 TCTGAGGGAAACAGTGAC 21
```

RESULT 8
AX431270
```
LOCUS       AX431270                1065 bp    DNA     linear   PAT 28-JUN-2002
DEFINITION  Sequence 4 from Patent WO0240650.
ACCESSION   AX431270
VERSION     AX431270.1  GI:21656152
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE     Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 4 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES             Location/Qualifiers
     source          1..1065
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for K2S protein"
BASE COUNT      225 a    314 c    312 g    214 t
ORIGIN
```

Query Match   100.0%; Score 18; DB 6; Length 1065;
  Best Local Similarity   100.0%; Pred. No. 1.1e+02;
  Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db          1 TCTGAGGGAAACAGTGAC 18
```

RESULT 9
AX431273
```
LOCUS       AX431273                1065 bp    DNA     linear   PAT 28-JUN-2002
DEFINITION  Sequence 7 from Patent WO0240650.
ACCESSION   AX431273
VERSION     AX431273.1  GI:21656155
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE     Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 7 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES             Location/Qualifiers
     source          1..1065
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for K2S protein"
BASE COUNT      225 a    314 c    312 g    214 t
ORIGIN
```

Query Match   100.0%; Score 18; DB 6; Length 1065;
  Best Local Similarity   100.0%; Pred. No. 1.1e+02;
  Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db          1 TCTGAGGGAAACAGTGAC 18
```

RESULT 10
E01943
```
LOCUS       E01943                  1065 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01943
VERSION     E01943.1  GI:2170191
KEYWORDS    JP 1989104167-A/10.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1065)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
```

```
TITLE     Kobayashi,M.
JOURNAL   NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
          Patent: JP 198910416T-A 10 21-APR-1989;
          FUJISAWA PHARMACEUT CO LTD
COMMENT   OS   Artificial gene
          OC   Artificial sequence; Genes.
          PN   JP 198910416T-A/10
          PD   21-APR-1989
          PF   01-AUG-1988 JP 1988192320
          PR   03-AUG-1987 GB 87  8718298, 26-OCT-1987 GB 87  8725052, PR
          13-NOV-1987 GB 87  8726683
          PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
          PI   NOTANI JOJI,
          PI   KOBAYASHI MASAKAZU
          PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64,(C12R1:19),
          PC   (C12N9/64,C12R1:91);
          CC   strandedness: Double;
          CC   topology: Linear;
          CC   hypothetical: No;
          CC   anti-sense: No;
          CC   *source: clone=pthTTtrp;
          FH   Key             Location/Qualifiers
          FH
          FT   CDS             1..1065
          FT                   /product='new t-PA'.
FEATURES          Location/Qualifiers
     source       1..1065
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
BASE COUNT      225 a    314 c    311 g    215 t
ORIGIN

Query Match        100.0%;  Score 18;  DB 6;  Length 1065;
Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGAC 21

RESULT 11
A27431
LOCUS       A27431                1068 bp     DNA      linear    PAT 09-JUL-2002
DEFINITION  DNA sequence of coding regionin pTTkPAdeltatrp.
ACCESSION   A27431
VERSION     A27431.1  GI:21727228
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 31 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES          Location/Qualifiers
     source       1..1068
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
BASE COUNT      226 a    314 c    313 g    215 t
ORIGIN

Query Match        100.0%;  Score 18;  DB 6;  Length 1068;
Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
```

```
Db    4 TCTGAGGGAAACAGTGAC 21

RESULT 12
A27433
LOCUS       A27433                1068 bp     DNA      linear    PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pTTiPAdeltatrp.
ACCESSION   A27433
VERSION     A27433.1  GI:2172T229
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 33 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES          Location/Qualifiers
     source       1..1068
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
BASE COUNT      225 a    315 c    312 g    216 t
ORIGIN

Query Match        100.0%;  Score 18;  DB 6;  Length 1068;
Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGAC 21

RESULT 13
A27439
LOCUS       A27439                1068 bp     DNA      linear    PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pSTTktrp.
ACCESSION   A27439
VERSION     A27439.1  GI:21727232
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 39 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES          Location/Qualifiers
     source       1..1068
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
BASE COUNT      228 a    313 c    312 g    215 t
ORIGIN

Query Match        100.0%;  Score 18;  DB 6;  Length 1068;
Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGAC 21

RESULT 14
A27445
LOCUS       A27445                1068 bp     DNA      linear    PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in puTTtrp.
```

```
ACCESSION    A27445
VERSION      A27445.1    GI:21727234
KEYWORDS      .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1  (bases 1 to 1068)
  AUTHORS    Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
             Kobayashi,M.
  TITLE      New tissue plasminogen activator
  JOURNAL    Patent: EP 0302456-A 45 08-FEB-1989;
             FUJISAWA PHARMACEUTICAL CO., LTD
  FEATURES   Location/Qualifiers
     source  1..1068
             /organism="synthetic construct"
             /mol_type="genomic DNA"
             /db_xref="taxon:32630"
BASE COUNT       226 a     314 c     313 g     215 t
ORIGIN

Query Match             100.0%;  Score 18;  DB 6;  Length 1068;
Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGAC 21


RESULT 15
AR059987
LOCUS        AR059987                  1068 bp    DNA          linear    PAT 29-SEP-1999
DEFINITION   Sequence 44 from patent US 5840533.
ACCESSION    AR059987
VERSION      AR059987.1   GI:5986437
KEYWORDS      .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 1068)
  AUTHORS    Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
             Kobayashi,M.
  TITLE      Tissue plasminogen activator
  JOURNAL    Patent: US 5840533-A 44 24-NOV-1998;
  FEATURES   Location/Qualifiers
     source  1..1068
             /organism="unknown"
BASE COUNT       226 a     314 c     313 g     215 t
ORIGIN

Query Match             100.0%;  Score 18;  DB 6;  Length 1068;
Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    4 TCTGAGGGAAACAGTGAC 21


Search completed: August 19, 2003, 04:46:24
Job time : 77.5794 secs
```

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    August 18, 2003, 19:51:53 ; Search time 6.34127 Seconds
                                        (without alignments)
                          7662.485 Million cell updates/sec

Title:           US-09-987-455-1
Perfect score:   18
Sequence:        1 tctcgaggaaacagtgac 18
Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters:    5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  N_Geneseq_19Jun03:*
   1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
   2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
   3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
   4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
   5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
   6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
   7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
   8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
   9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
  10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
  11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
  12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
  13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
  14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
  15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
  16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
  17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
  18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
  19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
  20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
  21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
  22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
  23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
  24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
  25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	18	100.0	18	24	AAD40613	OmpA N-terminal pe
2	18	100.0	18	24	AAD38989	Human K2S heterolo
c 3	18	100.0	42	17	AAT15496	Megakaryocyte diff
4	18	100.0	42	24	AAD40621	Human tPA gene amp
5	18	100.0	42	24	AAD38990	Human tissue plasm
6	18	100.0	55	17	AAT16011	Human tissue-plasm
c 7	18	100.0	952	9	AAN80896	Recombinant human
8	18	100.0	1047	14	AAQ40658	tPA-2 cDNA. Synth
9	18	100.0	1047	9	AAQ40664	tPA-8 cDNA. Synth
10	18	100.0	1065	24	AAD40616	Human kringle 2 se
11	18	100.0	1066	10	AAN91128	Sequence of coding
12	18	100.0	1068	10	AAN91120	Sequence of coding
13	18	100.0	1068	10	AAN91122	Sequence of coding
14	18	100.0	1068	10	AAN91124	Sequence of coding
15	18	100.0	1068	14	AAQ53320	Human truncated tP
16	18	100.0	1068	14	AAQ53322	Human truncated tP
17	18	100.0	1068	14	AAQ40657	tPA-1 cDNA. Synth
18	18	100.0	1068	14	AAQ40660	tPA-6 cDNA. Synth
19	18	100.0	1068	14	AAQ40663	tPA-7 cDNA. Synth
20	18	100.0	1068	14	AAQ40667	tPA-11 cDNA. Synt
21	18	100.0	1068	14	AAQ40669	tPA-12 cDNA. Synt
22	18	100.0	1068	14	AAQ40671	tPA-13 cDNA. Synt
23	18	100.0	1068	14	AAQ40673	tPA-14 cDNA. Synt
24	18	100.0	1068	14	AAQ40677	tPA-16 cDNA. Synt
25	18	100.0	1068	14	AAQ40679	tPA-17 cDNA. Synt
26	18	100.0	1068	14	AAQ40681	tPA-18 cDNA. Synt
27	18	100.0	1068	14	AAQ40683	tPA-5 cDNA. Synth
28	18	100.0	1068	14	AAQ40685	tPA-6 cDNA. Synth
29	18	100.0	1068	14	AAQ40675	tPA-15 cDNA. Synt
30	18	100.0	1125	10	AAN91129	Sequence of coding
31	18	100.0	1128	24	AAD40614	Sequence of coding
32	18	100.0	1170	10	AAN91135	Sequence of coding
33	18	100.0	1170	10	AAN91136	Sequence of coding
34	18	100.0	1314	10	AAN91117	Sequence of coding
35	18	100.0	1314	10	AAN91123	Sequence of coding
36	18	100.0	1314	10	AAN91126	Sequence of coding
37	18	100.0	1314	10	AAN91127	Sequence encoding
38	18	100.0	1341	11	AAQ06762	Sequence of coding
39	18	100.0	1419	10	AAN91133	Sequence of coding
40	18	100.0	1437	9	AAN81069	Mutant no. 2600: t
41	18	100.0	1440	9	AAN81492	Modified tissue pl
42	18	100.0	1440	9	AAN82183	Modified tissue pl
43	18	100.0	1440	19	AAV47648	Fibroblast growth
44	18	100.0	1440	20	AAZ32220	Human FGFR1-tPA fu
45	18	100.0	1440	20	AAX59330	cDNA coding for hu

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

```
RESULT 1
AAD40613
ID   AAD40613 standard; DNA; 18 BP.
XX
AC   AAD40613;
XX
DT   30-OCT-2002 (first entry)
XX
DE   OmpA N-terminal peptide encoding DNA #1.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..18
FT                   /*tag= a
FT                   /product= "OmpA N-terminal peptide"
FT                   /note= "No start and stop codon"
FT                   /partial
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
```

XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25033.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 2; Page 29; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 plus serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA fragment encoding OmpA N-terminal peptide.
XX
SQ   Sequence 18 BP; 6 A; 3 C; 6 G; 3 T; 0 other;

Query Match   100.0%; Score 18; DB 24; Length 18;
Best Local Similarity   100.0%; Pred. No. 10;
Matches   18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
               ||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGAC 18

RESULT 2
AAD38989
ID   AAD38989 standard; DNA; 18 BP.
XX
AC   AAD38989;
XX
DT   23-SEP-2002 (first entry)
XX
DE   Human K2S heterologous protein N-terminal peptide DNA #2.
XX
KW   Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;
KW   serine protease; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..12
FT   /*tag= a
FT   /product= "Human K2S heterologous protein N-terminal
FT   peptide #2"
XX
PN   WO200240696-A2.
XX
PD   23-MAY-2002.
XX
PF   08-NOV-2001; 2001WO-EP12920.
XX
PR   14-NOV-2000; 2000GB-0027782.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX

DR   WPI; 2002-471625/50.
XX
PT   Producing recombinant DNA-derived kringle 2 plus serine protease,
PT   comprises using a prokaryotic cell expressing a vector having a DNA
PT   coding for a heterologous protein operably linked to a DNA coding for
PT   the signal peptide OmpA -
XX
PS   Disclosure; Page 22; 52pp; English.
XX
CC   The invention relates to a method for producing recombinant DNA-derived
CC   heterologous protein in prokaryotic cells, where the heterologous protein
CC   is secreted extracellularly as an active and correctly folded protein and
CC   the prokaryotic cell contains and expresses a vector comprising the DNA
CC   coding for the heterologous protein operably linked to the DNA coding
CC   for the signal peptide OmpA or its functional derivative. The method is
CC   useful for commercial large-scale production of heterologous proteins,
CC   e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is
CC   generally applicable in the expression of several different proteins and
CC   polypeptides which do not require mammalian glycosylation in prokaryotic
CC   host cells. The method may also be used to obtain DNA sequences of a
CC   protein of interest to be expressed from databases and cloned for use.
CC   The present sequence is human K2S heterologous protein N-terminal peptide
CC   DNA.
XX
SQ   Sequence 18 BP; 6 A; 3 C; 6 G; 3 T; 0 other;

Query Match   100.0%; Score 18; DB 24; Length 18;
Best Local Similarity   100.0%; Pred. No. 10;
Matches   18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
               ||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGAC 18

RESULT 3
AAT15496/c
ID   AAT15496 standard; DNA; 42 BP.
XX
AC   AAT15496;
XX
DT   20-SEP-1996 (first entry)
XX
DE   Megakaryocyte differentiation/proliferation factor, NTPO cDNA primer.
XX
KW   Megakaryocyte; differentiation factor; proliferation; recombinant;
KW   thrombocytopenia; NTPO; polymerase chain reaction; ss.
XX
OS   Synthetic.
XX
PN   WO9603434-A1.
XX
PD   08-FEB-1996.
XX
PF   25-JUL-1995;   95WO-JP01476.
XX
PR   25-JUL-1994;   94JP-0172497.
XX
PA   (EISA) EISAI CO LTD.
XX
PI   Komatsu N, Okada M, Yoshitake S;
XX
DR   WPI; 1996-117001/12.
XX
PT   Megakaryocyte differentiation/proliferation factor - and fragments
PT   of it are prepared by recombinant techniques and used for treatment
PT   of thrombocytopenia.
XX
PS   Example 1; Page 12; 41pp; Japanese.
XX
CC   AAT15487-T15498 are PCR primers used for the isolation and amplification
CC   of cDNA encoding a human megakaryocyte differentiation/proliferation
CC   factor, NTPO, some of the primers may also be used to engineer

CC   recombinant NTPO DNA. DNA encoding the factor can be used to produce
CC   a recombinant NTPO at a high yield. NTPO is useful for the treatment
CC   of thrombocytopenia.
XX
SQ   Sequence 42 BP; 10 A; 9 C; 11 G; 12 T; 0 other;

Query Match   100.0%; Score 18; DB 17; Length 42;
Best Local Similarity   100.0%; Pred. No. 11;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
               ||||||||||||||||||
Db   31 TCTGAGGGAAACAGTGAC 14

RESULT 4
AAD40621
ID   AAD40621 standard; DNA; 42 BP.
XX
AC   AAD40621;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Human tPA gene amplifying PCR primer, sK2/174.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; PCR; primer; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Example 1; Page 80; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a PCR primer which is used for amplifying human tissue plasminogen
CC   activator (tPA) gene. This primer is used in the exemplification of
CC   the invention.
XX
SQ   Sequence 42 BP; 10 A; 9 C; 19 G; 4 T; 0 other;

Query Match   100.0%; Score 18; DB 24; Length 42;
Best Local Similarity   100.0%; Pred. No. 11;

Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
               ||||||||||||||||||
Db   25 TCTGAGGGAAACAGTGAC 42

RESULT 5
AAD38990
ID   AAD38990 standard; DNA; 42 BP.
XX
AC   AAD38990;
XX
DT   23-SEP-2002 (first entry)
XX
DE   Human tissue plasminogen activator (tPA) gene amplifying primer, sK2/174.
XX
KW   Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;
KW   serine protease; human; primer; PCR; ss.
XX
OS   Homo sapiens.
XX
PN   WO200240696-A2.
XX
PD   23-MAY-2002.
XX
PF   08-NOV-2001; 2001WO-EP12920.
XX
PR   14-NOV-2000; 2000GB-0027782.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-471625/50.
XX
PT   Producing recombinant DNA-derived kringle 2 plus serine protease,
PT   comprises using a prokaryotic cell expressing a vector having a DNA
PT   coding for a heterologous protein operably linked to a DNA coding for
PT   the signal peptide OmpA -
XX
PS   Example 1; Page 9; 52pp; English.
XX
CC   The invention relates to a method for producing recombinant DNA-derived
CC   heterologous protein in prokaryotic cells, where the heterologous protein
CC   is secreted extracellularly as an active and correctly folded protein and
CC   the prokaryotic cell contains and expresses a vector comprising the DNA
CC   coding for the heterologous protein operably linked to the DNA coding
CC   for the signal peptide OmpA or its functional derivative. The method is
CC   useful for commercial large-scale production of heterologous proteins,
CC   e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is
CC   generally applicable in the expression of several different proteins and
CC   polypeptides which do not require mammalian glycosylation in prokaryotic
CC   host cells. The method may also be used to obtain DNA sequences of a
CC   protein of interest to be expressed from databases and cloned for use.
CC   The present sequence is human tissue plasminogen activator (tPA) gene
CC   amplifying PCR primer.
XX
SQ   Sequence 42 BP; 10 A; 9 C; 19 G; 4 T; 0 other;

Query Match   100.0%; Score 18; DB 24; Length 42;
Best Local Similarity   100.0%; Pred. No. 11;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
               ||||||||||||||||||
Db   25 TCTGAGGGAAACAGTGAC 42

RESULT 6
AAT16011/c
ID   AAT16011 standard; DNA; 55 BP.
XX

```
AC   AAT16011;
XX
DT   23-SEP-1996 (first entry)
XX
DE   Human tissue-plasminogen activator (t-PA) gene N-terminal PCR primer.
XX
KW   Megakaryocyte; differentiation factor; proliferation; recombinant;
KW   thrombocytopenia; NTPO; polymerase chain reaction; ss.
XX
OS   Synthetic.
XX
PN   WO9603434-A1.
XX
PD   08-FEB-1996.
XX
PF   25-JUL-1995;   95WO-JP01476.
XX
PR   25-JUL-1994;   94JP-0172497.
XX
PA   (EISA ) EISAI CO LTD.
XX
PI   Komatsu N, Okada M, Yoshitake S;
XX
DR   WPI; 1996-117001/12.
XX
PT   Megakaryocyte differentiation/proliferation factor - and fragments
PT   of it are prepared by recombinant techniques and used for treatment
PT   of thrombocytopenia
XX
PS   Example 4; Page 13; 41pp; Japanese.
XX
CC   AAT16011-T16014 are PCR primers used for the isolation and amplification
CC   of the human tissue-plasminogen activator (t-PA) gene. The T-PA gene
CC   may be incorporated into a plasmid contg. DNA encoding a megakaryocyte
CC   differentiation/proliferation factor, NTPO, so as expression produces
CC   a chimeric t-PA-NTPO protein. DNA encoding the factor can be used to
CC   produce a recombinant NTPO at a high yield. NTPO is useful for the
CC   treatment of thrombocytopenia either alone or in conjunction with
CC   t-PA.
XX
SQ   Sequence 55 BP; 12 A; 11 C; 17 G; 15 T; 0 other;

Query Match             100.0%;  Score 18;  DB 17;  Length 55;
Best Local Similarity   100.0%;  Pred. No. 12;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
Db    40 TCTGAGGGAAACAGTGAC 23

RESULT 7
AAN80896
ID   AAN80896 standard; cDNA; 952 BP.
XX
AC   AAN80896;
XX
DT   31-OCT-1990 (first entry)
XX
DE   Recombinant human tissue plasminogen activator A chain sequence.
XX
KW   human tissue plasminogen activator (t-PA) A chain;
KW   antithrombotic agents; ds.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   sig_peptide    14..118
FT                  /*tag= a
FT   mat_peptide    119..943
FT                  /*tag= b
FT   misc_feature   915..952
FT                  /product=t-PA A-chain
FT                  /*tag= c
FT                  /label=synthetic oligomer
FT                  /note="capable of being transcribed into an mRNA
                        stop codon"
XX
PN   WO8808878-A.
XX
PD   17-NOV-1988.
XX
PF   12-MAY-1988;   88WO-US01624.
XX
PR   15-MAY-1987;   87US-0050950.
PR   12-MAY-1988;   88WO-US01624.
XX
PA   (INTE-) INTEG GENETICS INC.
XX
PI   Reddy VB;
XX
DR   WPI; 1988-338227/47.
DR   P-PSDB; AAP80770.
XX
PT   DNA coding for tissue plasminogen activator A chain -
PT   used as labelled A chain for antithrombotic drug screening
XX
PS   Disclosure; Page ?; p; English.
XX
CC   Recombinant sequence was obtained upon ligation of a 184bp
CC   DdeI-EcoRI fragment of p7BDT-10, a 730bp SalI-EcoRI fragment of
CC   p7BDT-10 encoding N-terminal portion of t-PA and a synthetic
CC   oligomer capable of being transcribed into an mRNA stop codon.
CC   The ligated product was cleaved with SalI to give the 952bp
CC   fragment shown. This fragment was cloned into the XhoI site of
CC   bovine papilloma virus DNA and used to transfect C127 mouse cells.
XX
SQ   Sequence 952 BP; 225 A; 266 C; 286 G; 175 T; 0 other;

Query Match             100.0%;  Score 18;  DB 9;  Length 952;
Best Local Similarity   100.0%;  Pred. No. 16;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
Db   638 TCTGAGGGAAACAGTGAC 655

RESULT 8
AAQ40658
ID   AAQ40658 standard; cDNA; 1047 BP.
XX
AC   AAQ40658;
XX
DT   06-AUG-1993 (first entry)
XX
DE   tPA-2 cDNA.
XX
KW   Blood; tissue plasminogen activator; tPA; mutein; stability; tPA-1;
KW   physiological; activity; tPA-2; pTB 1127; mutagenesis; plasmid;
KW   truncated; deletion; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   mutation       369..370
FT                  /*tag= a
FT                  /note="Position of deletion mutation"
XX
PN   JP05076361-A.
XX
PD   30-MAR-1993.
XX
PF   10-MAY-1991;   91JP-0105689.
XX
PR   10-MAY-1990;   90JP-0118710.
```

PR   25-DEC-1990;   90JP-0405848.
XX
PA   (TAKE) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-139567/17.
DR   P-PSDB; AAR35401.
XX
PT   Tissue plasminogen activator mutein - useful for treating
PT   myocardial infarction and cerebral thrombosis
XX
PS   Disclosure; Page 26-27; 92pp; Japanese.
XX
CC   The sequences given in AAQ40657-58 are cDNAs encoding tissue plasminogen
CC   activator (tPA) muteins, tPA-1 and tPA-2. These muteins are truncated
CC   versions of tPA (see AAQ40654) and tPA-2 further contains a deletion of
CC   amino acids corresponding to amino acids 296-302 of tPA. The plasmid
CC   pTB 1127 was treated with the synthetic oligomers given in AAQ40655-56
CC   in an in vitro mutagenesis system. The tPA muteins, tPA-1 and -2 have
CC   good stability in blood and good physiological activity.
XX
SQ   Sequence 1047 BP; 220 A; 307 C; 305 G; 215 T; 0 other;

Query Match   100.0%; Score 18; DB 14; Length 1047;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 9
AAQ40664
ID   AAQ40664 standard; cDNA; 1047 BP.
XX
AC   AAQ40664;
XX
DT   06-AUG-1993 (first entry)
XX
DE   tPA-8 cDNA.
XX
KW   Blood; tissue plasminogen activator; tPA; mutein; stability; tPA-7;
KW   physiological; activity; tPA-8; pTB 1277; mutagenesis; plasmid;
KW   truncated; deletion; mutation; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   mutation   369..370
FT   /*tag= a
FT   /note= "Position of deletion mutation"
FT   mutation   371..376
FT   /*tag= b
FT   /note= "GAGCGG>TATCAC"
XX
PN   JP05076361-A.
XX
PD   30-MAR-1993.
XX
PF   10-MAY-1991;   91JP-0105689.
XX
PR   10-MAY-1990;   90JP-0118710.
PR   25-DEC-1990;   90JP-0405848.
XX
PA   (TAKE) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-139567/17.
DR   P-PSDB; AAR35404.
XX
PT   Tissue plasminogen activator mutein - useful for treating
PT   myocardial infarction and cerebral thrombosis
XX
PS   Disclosure; Page 31-32; 92pp; Japanese.

XX
CC   The sequences given in AAQ40663-64 are cDNAs encoding tissue plasminogen
CC   activator (tPA) muteins, tPA-7 and tPA-8. These muteins are truncated
CC   versions of tPA (see AAQ40654). tPA-7 further comprises a Glu-Arg>
CC   Tyr-His mutation amino acids corresponding to amino acids 303-304 of
CC   tPA and tPA-8 further contains a deletion of amino acids corresponding
CC   to amino acids 296-302 of tPA, and the mutation contained in tPA-7.
CC   The plasmid pTB 1277 was treated with the synthetic oligomers given
CC   in AAQ40661-62 in an in vitro mutagenesis system. The tPA muteins,
CC   tPA-7 and -8 have good stability in blood and good physiological
CC   activity.
XX
SQ   Sequence 1047 BP; 221 A; 308 C; 301 G; 217 T; 0 other;

Query Match   100.0%; Score 18; DB 14; Length 1047;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 10
AAD40616
ID   AAD40616 standard; DNA; 1065 BP.
XX
AC   AAD40616;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Human kringle 2 serine protease (K2S) DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; enzyme; human; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1065
FT   /*tag= a
FT   /product= "Human K2S protein"
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001;   2001WO-EP12857.
XX
PR   14-NOV-2000;   2000GB-0027779.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25036.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 12; Page 31; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.

```
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is human K2S DNA.
XX
SQ   Sequence 1065 BP; 225 A; 314 C; 312 G; 214 T; 0 other;

Query Match        100.0%;  Score 18;  DB 24;  Length 1065;
Best Local Similarity   100.0%;  Pred. No. 17;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db        1 TCTGAGGGAAACAGTGAC 18

RESULT 11
AAN91128
ID   AAN91128 standard; DNA; 1066 BP.
XX
AC   AAN91128;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of coding region in plasmid puTTtrp.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key           Location/Qualifiers
FT   CDS           1..1065
FT                 /*tag=  a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M,  Saito Y,  Sasaki H,  Hayashi M,  Notani J,  Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94414.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1066 BP; 225 A; 314 C; 312 G; 215 T; 0 other;

Query Match        100.0%;  Score 18;  DB 10;  Length 1066;
Best Local Similarity   100.0%;  Pred. No. 17;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db        4 TCTGAGGGAAACAGTGAC 21

RESULT 12
AAN91120
ID   AAN91120 standard; DNA; 1068 BP.
XX
AC   AAN91120;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of coding region in plasmid pST112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key           Location/Qualifiers
FT   CDS           1..1065
FT                 /*tag=  a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M,  Saito Y,  Sasaki H,  Hayashi M,  Notani J,  Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94407.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1068 BP; 226 A; 314 C; 313 G; 215 T; 0 other;

Query Match        100.0%;  Score 18;  DB 10;  Length 1068;
Best Local Similarity   100.0%;  Pred. No. 17;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db        4 TCTGAGGGAAACAGTGAC 21

RESULT 13
AAN91122
ID   AAN91122 standard; DNA; 1068 BP.
XX
AC   AAN91122;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of coding region in plasmid pTTiPA delta trp.
```

```
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1065
FT                   /tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94408.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1068 BP; 225 A; 315 C; 312 G; 216 T; 0 other;

Query Match           100.0%;   Score 18;   DB 10;   Length 1068;
Best Local Similarity 100.0%;   Pred. No. 17;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1   TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
Db   4   TCTGAGGGAAACAGTGAC 21

RESULT 14
AAN91124
ID   AAN91124 standard; DNA; 1068 BP.
XX
AC   AAN91124;
XX
DT   25-MAR-2003   (updated)
DT   03-OCT-2002   (updated)
DT   18-JUN-1990   (first entry)
XX
DE   Sequence of coding region in plasmid pSTTktrp.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1065
FT                   /tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
```

```
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94411.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1068 BP; 228 A; 313 C; 312 G; 215 T; 0 other;

Query Match           100.0%;   Score 18;   DB 10;   Length 1068;
Best Local Similarity 100.0%;   Pred. No. 17;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1   TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
Db   4   TCTGAGGGAAACAGTGAC 21

RESULT 15
AAQ53320
ID   AAQ53320 standard; cDNA to mRNA; 1068 BP.
XX
AC   AAQ53320;
XX
DT   20-JUN-1994   (first entry)
XX
DE   Human truncated tPA-1 mutein coding sequence.
XX
KW   Tissue plasminogen activator; t-PA; mutein; fibrin; antigen;
KW   anti-fibrin; monoclonal antibody; hybridoma; thrombolysis
KW   targetted antithrombotic agent; hybridisation probe;
KW   Kringle domain; protease; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1068
FT                   /tag= a
FT                   /product= tPA-1
FT                   /note= "encodes Kringle-2 and protease domains"
XX
PN   JP05304992-A.
XX
PD   19-NOV-1993.
XX
PF   17-JUN-1992;   92JP-0158301.
XX
PR   20-JUN-1991;   91JP-0148936.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-408334/51.
DR   P-PSDB; AAR44836.
XX
PT   Hybrid monoclonal antibody - used for prepn. of thrombolytic drug
PT   having increased thrombolytic activity and specificity and
PT   reduced reactivity to fibrinogen
XX
PS   Disclosure; Fig 7; 38pp; Japanese.
```

```
XX
CC      Site-directed mutagenesis was carried out on a human tPA cDNA
CC      template. The resulting mutated coding sequence encoded a Met
CC      residue followed by amino acids 174-527 (i.e. the Kringle 2
CC      and protease domains) of mature wild-type tPA. The mutein has
CC      increased thrombolytic activity and very little reactivity to
CC      fibrinogen. Monoclonal antibodies specific for fibrin are used in
CC      the production of bispecific monoclonal antibodies which also
CC      recognise truncated tPA muteins such as tPA-1 lacking the finger,
CC      EGF and Kringle 1 domains.
XX
SQ      Sequence 1068 BP; 226 A; 314 C; 313 G; 215 T; 0 other;

Query Match       100.0%; Score 18; DB 14; Length 1068;
Best Local Similarity   100.0%; Pred. No. 17;
Matches   18; Conservative  0; Mismatches   0; Indels   0; Gaps   0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db      4 TCTGAGGGAAACAGTGAC 21


Search completed: August 18, 2003, 23:28:45
Job time : 6.34127 secs
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003   Compugen Ltd.

Run on: August 18, 2003, 22:51:23 ; Search time 1.40873 Seconds
                         (without alignments)
                         5639.757 Million cell updates/sec

Title: US-09-987-455-1
Perfect score: 18
Sequence: 1 tctgagggaaacagtgac 18

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents NA:*
             1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
             2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
             3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
             4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
             5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
             6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Query Match	Length	DB	ID	Description
1		18	100.0	18	1065	2	US-08-811-949-60	Sequence 60, Appl
2		18	100.0	18	1068	2	US-08-811-949-44	Sequence 44, Appl
3		18	100.0	18	1068	2	US-08-811-949-46	Sequence 46, Appl
4		18	100.0	18	1068	2	US-08-811-949-52	Sequence 52, Appl
5		18	100.0	18	1068	2	US-08-811-949-58	Sequence 58, Appl
6		18	100.0	18	1170	2	US-08-811-949-64	Sequence 64, Appl
7		18	100.0	18	1170	2	US-08-811-949-66	Sequence 66, Appl
8		18	100.0	18	1314	2	US-08-811-949-48	Sequence 48, Appl
9		18	100.0	18	1314	2	US-08-811-949-50	Sequence 50, Appl
10		18	100.0	18	1314	2	US-08-811-949-54	Sequence 54, Appl
11		18	100.0	18	1314	2	US-08-811-949-56	Sequence 56, Appl
12		18	100.0	18	1419	2	US-08-811-949-62	Sequence 62, Appl
13		18	100.0	18	1724	6	5200340-5	Patent No. 5200340
14		18	100.0	18	1738	6	5200340-1	Patent No. 5200340
15		18	100.0	18	1848	3	US-08-814-412-10	Sequence 10, Appl
16		18	100.0	18	1955	2	US-08-883-795A-39	Sequence 39, Appl
17		18	100.0	18	1974	2	US-08-811-949-38	Sequence 38, Appl
18		18	100.0	18	2101	2	US-08-811-949-42	Sequence 42, Appl
19		18	100.0	18	2162	1	US-08-119-512-3	Sequence 3, Appli
20		18	100.0	18	2162	1	US-08-488-015B-3	Sequence 3, Appli
21		18	100.0	18	2162	1	US-08-488-015B-25	Sequence 25, Appl
22		18	100.0	18	2457	6	5344773-1	Patent No. 5344773
23		18	100.0	18	2497	6	5185259-2	Patent No. 5185259
24		18	100.0	18	7360	1	US-08-286-740-1	Sequence 1, Appli
25		18	100.0	18	7360	5	PCT-US95-09576-1	Sequence 1, Appli
26		15.4	85.6		1459	3	US-09-020-956-174	Sequence 174, App
27		15.4	85.6		1459	3	US-09-030-607-174	Sequence 174, App
28		15.4	85.6		1459	4	US-09-439-313-174	Sequence 174, App
29		15.4	85.6		1459	4	US-09-352-616A-174	Sequence 174, App
30		15.4	85.6		1459	4	US-09-232-149A-174	Sequence 174, App
31		15.4	85.6		3772	4	US-09-996-243-61	Sequence 61, Appli
32		15	83.3		5159	2	US-08-146-930-3	Sequence 3, Appli
33	c	15	83.3		5159	5	US-08-458-240-3	Sequence 3, Appli
34	c	15	83.3		5159	5	PCT-US93-03993-3	Sequence 3, Appli
35	c	15	83.3		35100	5	US-08-306-691B-19	Sequence 19, Appl
36	c	15	83.3		35100	5	PCT-US93-06251-19	Sequence 19, Appl
37		14.8	82.2		20	3	US-08-930-601-6	Sequence 6, Appli
38		14.8	82.2		3222	4	US-09-206-942-48	Sequence 48, Appl
39		14.8	82.2		3240	4	US-09-206-942-46	Sequence 46, Appl
40		14.8	82.2		64467	4	US-09-803-671B-3	Sequence 3, Appli
41	c	14.4	80.0		1881	5	PCT-US94-09752-2	Sequence 2, Appli
42		14.4	80.0		6908	4	US-09-620-312D-977	Sequence 977, App
43		14	77.8		34185	4	US-09-545-481-3	Sequence 3, Appli
44	c	13.8	76.7		89	2	US-08-373-190-51	Sequence 51, Appl
45	c	13.8	76.7		89	2	US-08-438-190A-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-08-811-949-60
; Sequence 60, Application US/0811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
;   INFORMATION FOR SEQ ID NO: 60:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1065 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1062
US-08-811-949-60

Query Match   100.0%; Score 18; DB 2; Length 1065;

Best Local Similarity 100.0%;   Pred. No. 2.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 2
US-08-811-949-44
; Sequence 44, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 44:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1065
US-08-811-949-44

Query Match   100.0%; Score 18; DB 2; Length 1068;
Best Local Similarity   100.0%; Pred. No. 2.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 3
US-08-811-949-46
; Sequence 46, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 46:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1065
US-08-811-949-46

Query Match   100.0%; Score 18; DB 2; Length 1068;
Best Local Similarity   100.0%; Pred. No. 2.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 4
US-08-811-949-52
; Sequence 52, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:

;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1065
US-08-811-949-52

Query Match   100.0%;   Score 18;   DB 2;   Length 1068;
Best Local Similarity 100.0%;   Pred. No. 2.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 5
US-08-811-949-58
; Sequence 58, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1065
US-08-811-949-58

Query Match   100.0%;   Score 18;   DB 2;   Length 1068;
Best Local Similarity 100.0%;   Pred. No. 2.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGAC 21

RESULT 6
US-08-811-949-64
; Sequence 64, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1170 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1167
US-08-811-949-64

Query Match   100.0%;   Score 18;   DB 2;   Length 1170;
Best Local Similarity 100.0%;   Pred. No. 2.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TCTGAGGGAAAACAGTGAC 18
        |||||||||||||||||||
Db  106 TCTGAGGGAAAACAGTGAC 123
```

RESULT 7
US-08-811-949-66
; Sequence 66, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 66:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1170 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1167
US-08-811-949-66

```
Query Match         100.0%;  Score 18;  DB 2;  Length 1170;
Best Local Similarity  100.0%;  Pred. No. 2.9;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TCTGAGGGAAAACAGTGAC 18
        |||||||||||||||||||
Db  106 TCTGAGGGAAAACAGTGAC 123
```

RESULT 8
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO

;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 48:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1314 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1311
US-08-811-949-48

```
Query Match         100.0%;  Score 18;  DB 2;  Length 1314;
Best Local Similarity  100.0%;  Pred. No. 2.9;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TCTGAGGGAAAACAGTGAC 18
        |||||||||||||||||||
Db  250 TCTGAGGGAAAACAGTGAC 267
```

RESULT 9
US-08-811-949-50
; Sequence 50, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible

```
         OPERATING SYSTEM: PC-DOS/MS-DOS
         SOFTWARE: PatentIn Release #1.0, Version #1.30
    ; CURRENT APPLICATION DATA:
    ;   APPLICATION NUMBER: US/08/811,949
    ;   FILING DATE: 05-MAR-1997
    ;   CLASSIFICATION: 435
    ; ATTORNEY/AGENT INFORMATION:
    ;   NAME: OBLON, NORMAN F.
    ;   REGISTRATION NUMBER: 24,618
    ;   REFERENCE/DOCKET NUMBER: 18-966-0
    ; TELECOMMUNICATION INFORMATION:
    ;   TELEPHONE: 703-413-3000
    ;   TELEFAX: 703-413-2220
    ; INFORMATION FOR SEQ ID NO: 50:
    ;   SEQUENCE CHARACTERISTICS:
    ;     LENGTH: 1314 base pairs
    ;     TYPE: nucleic acid
    ;     STRANDEDNESS: double
    ;     TOPOLOGY: circular
    ;   MOLECULE TYPE: DNA (genomic)
    ;   FEATURE:
    ;     NAME/KEY: CDS
    ;     LOCATION: 1..1311
    US-08-811-949-50

    Query Match             100.0%;  Score 18;  DB 2;  Length 1314;
    Best Local Similarity   100.0%;  Pred. No. 2.9;
    Matches 18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

    Qy      1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
    Db    250 TCTGAGGGAAACAGTGAC 267


    RESULT 10
    US-08-811-949-54
    ; Sequence 54, Application US/0811949
    ; Patent No. 5840533
    ; GENERAL INFORMATION:
    ;   APPLICANT: NIWA, MINEO
    ;   APPLICANT: SAITO, YOSHIMASA
    ;   APPLICANT: SASAKI, HITOSHI
    ;   APPLICANT: HAYASHI, MASAKO
    ;   APPLICANT: NOTANI, JOUJI
    ;   APPLICANT: KOBAYASHI, MASAKAZU
    ;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
    ;   NUMBER OF SEQUENCES: 67
    ;   CORRESPONDENCE ADDRESS:
    ;     ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
    ;     ADDRESSEE: P.C.
    ;     STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
    ;     CITY: ARLINGTON
    ;     STATE: VA
    ;     COUNTRY: USA
    ;     ZIP: 22202
    ;   COMPUTER READABLE FORM:
    ;     MEDIUM TYPE: Floppy disk
    ;     COMPUTER: IBM PC compatible
    ;     OPERATING SYSTEM: PC-DOS/MS-DOS
    ;     SOFTWARE: PatentIn Release #1.0, Version #1.30
    ;   CURRENT APPLICATION DATA:
    ;     APPLICATION NUMBER: US/08/811,949
    ;     FILING DATE: 05-MAR-1997
    ;     CLASSIFICATION: 435
    ;   ATTORNEY/AGENT INFORMATION:
    ;     NAME: OBLON, NORMAN F.
    ;     REGISTRATION NUMBER: 24,618
    ;     REFERENCE/DOCKET NUMBER: 18-966-0
    ;   TELECOMMUNICATION INFORMATION:
    ;     TELEPHONE: 703-413-3000
    ;     TELEFAX: 703-413-2220
    ; INFORMATION FOR SEQ ID NO: 54:
    ;   SEQUENCE CHARACTERISTICS:
    ;     LENGTH: 1314 base pairs
    ;     TYPE: nucleic acid
    ;     STRANDEDNESS: double
    ;     TOPOLOGY: circular
    ;   MOLECULE TYPE: DNA (genomic)
    ;   FEATURE:
    ;     NAME/KEY: CDS
    ;     LOCATION: 1..1311
    US-08-811-949-54

    Query Match             100.0%;  Score 18;  DB 2;  Length 1314;
    Best Local Similarity   100.0%;  Pred. No. 2.9;
    Matches 18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

    Qy      1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
    Db    250 TCTGAGGGAAACAGTGAC 267


    RESULT 11
    US-08-811-949-56
    ; Sequence 56, Application US/08811949
    ; Patent No. 5840533
    ; GENERAL INFORMATION:
    ;   APPLICANT: NIWA, MINEO
    ;   APPLICANT: SAITO, YOSHIMASA
    ;   APPLICANT: SASAKI, HITOSHI
    ;   APPLICANT: HAYASHI, MASAKO
    ;   APPLICANT: NOTANI, JOUJI
    ;   APPLICANT: KOBAYASHI, MASAKAZU
    ;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
    ;   NUMBER OF SEQUENCES: 67
    ;   CORRESPONDENCE ADDRESS:
    ;     ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
    ;     ADDRESSEE: P.C.
    ;     STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
    ;     CITY: ARLINGTON
    ;     STATE: VA
    ;     COUNTRY: USA
    ;     ZIP: 22202
    ;   COMPUTER READABLE FORM:
    ;     MEDIUM TYPE: Floppy disk
    ;     COMPUTER: IBM PC compatible
    ;     OPERATING SYSTEM: PC-DOS/MS-DOS
    ;     SOFTWARE: PatentIn Release #1.0, Version #1.30
    ;   CURRENT APPLICATION DATA:
    ;     APPLICATION NUMBER: US/08/811,949
    ;     FILING DATE: 05-MAR-1997
    ;     CLASSIFICATION: 435
    ;   ATTORNEY/AGENT INFORMATION:
    ;     NAME: OBLON, NORMAN F.
    ;     REGISTRATION NUMBER: 24,618
    ;     REFERENCE/DOCKET NUMBER: 18-966-0
    ;   TELECOMMUNICATION INFORMATION:
    ;     TELEPHONE: 703-413-3000
    ;     TELEFAX: 703-413-2220
    ; INFORMATION FOR SEQ ID NO: 56:
    ;   SEQUENCE CHARACTERISTICS:
    ;     LENGTH: 1314 base pairs
    ;     TYPE: nucleic acid
    ;     STRANDEDNESS: double
    ;     TOPOLOGY: circular
    ;   MOLECULE TYPE: DNA (genomic)
    ;   FEATURE:
    ;     NAME/KEY: CDS
    ;     LOCATION: 1..1311
    US-08-811-949-56

    Query Match             100.0%;  Score 18;  DB 2;  Length 1314;
    Best Local Similarity   100.0%;  Pred. No. 2.9;
    Matches 18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

    Qy      1 TCTGAGGGAAACAGTGAC 18
```

```
Db       250 TCTGAGGGAAACAGTGAC 267
             |||||||||||||||||||
```

RESULT 12
US-08-811-949-62
; Sequence 62, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 62:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1419 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1416
US-08-811-949-62

Query Match   100.0%;   Score 18; DB 2; Length 1419;
Best Local Similarity 100.0%;   Pred. No. 3;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db   355 TCTGAGGGAAACAGTGAC 372

RESULT 13
5200340-5
;Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;   NUMBER OF SEQUENCES: 34
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/53,412
;   FILING DATE: 22-MAY-1987

;SEQ ID NO:5:
;   LENGTH: 1724
5200340-5

Query Match   100.0%;   Score 18; DB 6; Length 1724;
Best Local Similarity 100.0%;   Pred. No. 3;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db   640 TCTGAGGGAAACAGTGAC 657

RESULT 14
5200340-1
;Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;   NUMBER OF SEQUENCES: 34
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/53,412
;   FILING DATE: 22-MAY-1987
;SEQ ID NO:11:
;   LENGTH: 1738
5200340-1

Query Match   100.0%;   Score 18; DB 6; Length 1738;
Best Local Similarity 100.0%;   Pred. No. 3;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
            ||||||||||||||||||
Db   640 TCTGAGGGAAACAGTGAC 657

RESULT 15
US-08-814-412-10
; Sequence 10, Application US/08814412
; Patent No. 6150141
; GENERAL INFORMATION:
;   APPLICANT: Jarrell Ph.D., Kevin A.
;   TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
;   NUMBER OF SEQUENCES: 46
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Choate, Hall & Stewart
;   STREET: 53 State Street
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/814,412
;   FILING DATE: 11-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Jarrell Ph.D., Brenda H.
;   REGISTRATION NUMBER: 39,223
;   REFERENCE/DOCKET NUMBER: 0079571-0040
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617 248 5000
;   TELEFAX: 617 248 4000
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1848 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double

```
;    TOPOLOGY: linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;    CLONE:  amplified t-PA clone
US-08-814-412-10

Query Match       100.0%;  Score 18;  DB 3;  Length 1848;
Best Local Similarity  100.0%;  Pred. No. 3;
Matches   18;  Conservative  0;  Mismatches   0;  Indels  0;  Gaps  0;

Qy       1 TCTGAGGGAAACAGTGAC 18
           ||||||||||||||||||
Db     696 TCTGAGGGAAACAGTGAC 713

Search completed: August 19, 2003, 08:35:04
Job time : 4.40873 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 5.07937 Seconds
(without alignments)
7930.701 Million cell updates/sec

Title: US-09-987-455-1
Perfect score: 18
Sequence: 1 tctgagggaaacagtgac 18
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 1504479 seqs, 1118970152 residues
Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	100.0	18	11	US-09-987-457-5	Sequence 5, Appli
2	18	100.0	18	11	US-09-987-455-1	Sequence 1, Appli
3	18	100.0	42	11	US-09-987-457-6	Sequence 6, Appli
4	18	100.0	42	11	US-09-987-455-22	Sequence 22, Appli
5	18	100.0	1065	11	US-09-987-455-4	Sequence 4, Appli
6	18	100.0	1065	11	US-09-987-455-7	Sequence 7, Appli
7	18	100.0	1128	11	US-09-987-455-2	Sequence 2, Appli
8	18	100.0	1128	11	US-09-987-455-5	Sequence 5, Appli
9	18	100.0	1415	14	US-10-198-846-12748	Sequence 12748, A
10	18	100.0	1689	9	US-09-969-271-6	Sequence 6, Appli
11	18	100.0	2509	14	US-10-193-656-7	Sequence 7, Appli
12	18	100.0	2519	9	US-09-969-271-5	Sequence 5, Appli
13	18	100.0	2641	10	US-09-974-298-144	Sequence 144, App
14	15.4	85.6	139	10	US-09-878-574-237	Sequence 237, App
c 15	15.4	85.6	292	9	US-09-864-761-18405	Sequence 18405, A
c 16	15.4	85.6	352	10	US-09-960-352-6540	Sequence 6540, Ap
17	15.4	85.6	403	10	US-09-983-965-2988	Sequence 2988, Ap
18	15.4	85.6	403	10	US-09-983-965-4942	Sequence 4942, Ap
c 19	15.4	85.6	466	9	US-09-864-761-1647	Sequence 1647, Ap
c 20	15.4	85.6	530	13	US-10-027-632-39133	Sequence 39113, A
21	15.4	85.6	676	13	US-10-027-632-243054	Sequence 243054,
c 22	15.4	85.6	767	13	US-10-027-632-157405	Sequence 157405,
c 23	15.4	85.6	818	13	US-10-027-632-166380	Sequence 166380,
24	15.4	85.6	936	13	US-10-027-632-166215	Sequence 166215,
25	15.4	85.6	936	13	US-10-027-632-166216	Sequence 166216,
c 26	15.4	85.6	1184	13	US-10-027-632-101376	Sequence 101376,
27	15.4	85.6	1225	13	US-10-027-632-85099	Sequence 85099, A
c 28	15.4	85.6	1360	12	US-10-017-161-947	Sequence 947, App
29	15.4	85.6	1459	9	US-09-759-143-174	Sequence 174, App
30	15.4	85.6	1459	9	US-09-780-669-174	Sequence 174, App
31	15.4	85.6	1459	9	US-09-030-606-174	Sequence 174, App
32	15.4	85.6	1459	9	US-09-822-827-174	Sequence 174, App
33	15.4	85.6	1459	9	US-09-115-453-174	Sequence 174, App
34	15.4	85.6	1459	10	US-09-232-880-174	Sequence 174, App
35	15.4	85.6	1459	10	US-09-895-793-174	Sequence 174, App
36	15.4	85.6	1459	10	US-09-895-814-174	Sequence 174, App
37	15.4	85.6	1459	13	US-10-012-896-174	Sequence 174, App
38	15.4	85.6	1459	14	US-10-010-940-174	Sequence 174, App
39	15.4	85.6	1590	9	US-09-888-615-3	Sequence 3, Appli
c 40	15.4	85.6	1617	10	US-09-917-800A-1699	Sequence 1699, Ap
c 41	15.4	85.6	1617	12	US-10-205-194-175	Sequence 175, App
c 42	15.4	85.6	1901	13	US-10-027-632-266063	Sequence 266063,
43	15.4	85.6	2392	12	US-09-813-432-23	Sequence 23, Appl
44	15.4	85.6	3772	9	US-09-989-722-61	Sequence 61, Appl
45	15.4	85.6	3772	9	US-09-989-723-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1
US-09-987-457-5
; Sequence 5, Application US/09987457
; Publication No. US2003001315OA1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence of the N-terminal part of K2S molecule
US-09-987-457-5

Query Match 100.0%;   Score 18; DB 11; Length 18;
Best Local Similarity 100.0%;   Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGAC 18

RESULT 2

US-09-987-455-1
; Sequence 1, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence of N-terminal part of K2S protein
US-09-987-455-1

Query Match   100.0%; Score 18; DB 11; Length 18;
Best Local Similarity   100.0%; Pred. No. 7.1;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGAC 18

RESULT 3
US-09-987-457-6
; Sequence 6, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 6
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide
; OTHER INFORMATION: sequence
US-09-987-457-6

Query Match   100.0%; Score 18; DB 11; Length 42;
Best Local Similarity   100.0%; Pred. No. 7.2;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||

Db   25 TCTGAGGGAAACAGTGAC 42

RESULT 4
US-09-987-455-22
; Sequence 22, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 22
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-987-455-22

Query Match   100.0%; Score 18; DB 11; Length 42;
Best Local Similarity   100.0%; Pred. No. 7.2;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   25 TCTGAGGGAAACAGTGAC 42

RESULT 5
US-09-987-455-4
; Sequence 4, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-4

Query Match   100.0%; Score 18; DB 11; Length 1065;
Best Local Similarity   100.0%; Pred. No. 7.6;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    1 TCTGAGGGAAACAGTGAC 18
```

RESULT 6
US-09-987-455-7
; Sequence 7, Application US/09987455
; Publication No. US2003004972 9A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-7

Query Match 100.0%; Score 18; DB 11; Length 1065;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```
Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db    1 TCTGAGGGAAACAGTGAC 18
```

RESULT 7
US-09-987-455-2
; Sequence 2, Application US/09987455
; Publication No. US2003004972 9A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding

; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match 100.0%; Score 18; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```
Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db   67 TCTGAGGGAAACAGTGAC 84
```

RESULT 8
US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US2003004972 9A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match 100.0%; Score 18; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```
Qy    1 TCTGAGGGAAACAGTGAC 18
        ||||||||||||||||||
Db   67 TCTGAGGGAAACAGTGAC 84
```

RESULT 9
US-10-198-846-12748
; Sequence 12748, Application US/10198846
; Publication No. US2003009974 9A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12748
; LENGTH: 1415
; TYPE: DNA

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2, 3, 1414, 1415
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12748

Query Match             100.0%;  Score 18;  DB 14;  Length 1415;
Best Local Similarity   100.0%;  Pred. No. 7.7;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db        906 TCTGAGGGAAACAGTGAC 923


RESULT 10
US-09-969-271-6
;   Sequence 6, Application US/09969271
;   Patent No. US20020098179A1
;   GENERAL INFORMATION:
;   APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
;   APPLICANT: Pfizer Limited (GB and EP (GB) only)
;   TITLE OF INVENTION: Pharmaceutical Combinations
;   FILE REFERENCE: PCS10951APME
;   CURRENT APPLICATION NUMBER: US/09/969,271
;   CURRENT FILING DATE: 2001-10-01
;   PRIOR APPLICATION NUMBER: GB 0025473.0
;   PRIOR FILING DATE: 2000-10-17
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 6
;     LENGTH: 1689
;     TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-271-6

Query Match             100.0%;  Score 18;  DB 9;  Length 1689;
Best Local Similarity   100.0%;  Pred. No. 7.7;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db        625 TCTGAGGGAAACAGTGAC 642


RESULT 11
US-10-193-656-7
;   Sequence 7, Application US/10193656
;   Publication No. US2003009673A1
;   GENERAL INFORMATION:
;   APPLICANT: NY, Tor
;   APPLICANT: HOLMDAHL, Rikard
;   APPLICANT: LI, Jinan
;   TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
;   FILE REFERENCE: 3810/11577-US3
;   CURRENT APPLICATION NUMBER: US/10/193,656
;   CURRENT FILING DATE: 2002-07-10
;   PRIOR APPLICATION NUMBER: US 60/304,461
;   PRIOR FILING DATE: 2001-07-10
;   PRIOR APPLICATION NUMBER: US 60/304,490
;   PRIOR FILING DATE: 2001-07-10
;   PRIOR APPLICATION NUMBER: US 60/305,182
;   PRIOR FILING DATE: 2001-07-13
;   NUMBER OF SEQ ID NOS: 18
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 7
;     LENGTH: 2509
;     TYPE: DNA
;   ORGANISM: Homo sapiens
;   PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: GenBank / X07393
;   DATABASE ENTRY DATE: 1995-03-27
```

```
;   RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match             100.0%;  Score 18;  DB 14;  Length 2509;
Best Local Similarity   100.0%;  Pred. No. 7.7;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db        678 TCTGAGGGAAACAGTGAC 695


RESULT 12
US-09-969-271-5
;   Sequence 5, Application US/09969271
;   Patent No. US20020098179A1
;   GENERAL INFORMATION:
;   APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
;   APPLICANT: Pfizer Limited (GB and EP (GB) only)
;   TITLE OF INVENTION: Pharmaceutical Combinations
;   FILE REFERENCE: PCS10951APME
;   CURRENT APPLICATION NUMBER: US/09/969,271
;   CURRENT FILING DATE: 2001-10-01
;   PRIOR APPLICATION NUMBER: GB 0025473.0
;   PRIOR FILING DATE: 2000-10-17
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 5
;     LENGTH: 2519
;     TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-271-5

Query Match             100.0%;  Score 18;  DB 9;  Length 2519;
Best Local Similarity   100.0%;  Pred. No. 7.7;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db        701 TCTGAGGGAAACAGTGAC 718


RESULT 13
US-09-974-298-144
;   Sequence 144, Application US/09974298
;   Patent No. US20020156263A1
;   GENERAL INFORMATION:
;   APPLICANT: Chen, Huei-Mei
;   TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
;   FILE REFERENCE: PA-0037 P
;   CURRENT APPLICATION NUMBER: US/09/974,298
;   CURRENT FILING DATE: 2001-10-04
;   PRIOR APPLICATION NUMBER: 60/238,331
;   PRIOR FILING DATE: 2000-05-10
;   NUMBER OF SEQ ID NOS: 194
;   SOFTWARE: PERL Program
;   SEQ ID NO 144
;     LENGTH: 2641
;     TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   OTHER INFORMATION: Incyte ID No. US2002015263A1 1001470CB1
;   NAME/KEY: unsure
;   LOCATION: 2635
;   OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

Query Match             100.0%;  Score 18;  DB 10;  Length 2641;
Best Local Similarity   100.0%;  Pred. No. 7.7;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCTGAGGGAAACAGTGAC 18
```

```
Db          821 TCTGAGGGAAACAGTGAC 838
                |||||||||||||||||||

RESULT 14
US-09-878-574-237
; Sequence 237, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT: Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION: Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 237
;  LENGTH: 139
;  TYPE: DNA
;  ORGANISM: Glycine max
;  OTHER INFORMATION: Clone ID: LIB3028-053-Q1-B1-E8
US-09-878-574-237

Query Match             85.6%;  Score 15.4;  DB 10;  Length 139;
Best Local Similarity   94.1%;  Pred. No. 1.6e+02;
Matches    16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCTGAGGGAAACAGTGA 17
              ||||||||| |||||||
Db         74 TCTGAGGGAAAACAGTGA 90


RESULT 15
US-09-864-761-18405/c
; Sequence 18405, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT: Rank, David R.
;  APPLICANT: Hanzel, David K.
;  APPLICANT: Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663

;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18405
;  LENGTH: 292
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: MAP TO AP000459.2
;  OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 20
;  OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1e+02
;  OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 60
;  OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.5
;  OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 25
;  OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 14
;  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2
;  OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 61
;  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 23
;  OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 50
;  OTHER INFORMATION: EST HUMAN HIT: BE735581.1, EVALUE 7.00e-82
;  OTHER INFORMATION: SWISSPROT HIT: P02552, EVALUE 6.00e-41
;  OTHER INFORMATION: NT HIT: AF081484.1, EVALUE 4.00e-82
US-09-864-761-18405

Query Match             85.6%;  Score 15.4;  DB 9;  Length 292;
Best Local Similarity   94.1%;  Pred. No. 1.6e+02;
Matches    16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCTGAGGGAAACAGTGA 17
              ||||||||| |||||||
Db         77 TCTGAGGGAAGCAGTGA 61


Search completed: August 19, 2003, 14:22:19
Job time : 6.07937 secs
```

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 22:44:33 ; Search time 52.6587 Seconds
(without alignments)
8307.845 Million cell updates/sec

Title: US-09-987-455-1
Perfect score: 18
Sequence: 1 tctgaggaaacagtgac 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	18	100.0	241	9	AA376531	AA376531 EST88963
2	18	100.0	335	10	BE158218	BE158218 MR2-HT038
3	18	100.0	374	10	BE173847	BE173847 CMO-HT056
4	18	100.0	378	10	BE173842	BE173842 CMO-HT056
5	18	100.0	476	10	BE259860	601154266
6	18	100.0	519	14	CB149125	K-EST0205
7	18	100.0	582	10	BE386276	601273495
8	18	100.0	612	10	BE386111	601277221
9	18	100.0	639	9	AU117078	AU117078
10	18	100.0	644	10	BG477932	602522911
11	18	100.0	682	10	BE276291	601144209
12	18	100.0	690	10	BE895982	601438641
13	18	100.0	711	10	BG476511	602522049
14	18	100.0	725	12	BM042154	603615791
15	18	100.0	761	12	BI771956	603058981
16	18	100.0	778	10	BE408133	601303133
17	18	100.0	778	12	BG762705	602734613
18	18	100.0	798	10	BG281523	602402058
19	18	100.0	811	12	BI460599	603201239
20	18	100.0	818	9	AU124602	AU124602
21	18	100.0	831	13	BU156867	AGENCOURT
22	18	100.0	835	10	BG281371	602401866
23	18	100.0	841	13	BU157242	AGENCOURT
24	18	100.0	843	9	AU135634	AU135634
25	18	100.0	843	13	BQ084473	AGENCOURT
26	18	100.0	845	13	BQ077849	AGENCOURT
27	18	100.0	856	10	BG743550	602635588
28	18	100.0	858	10	BG281843	602403069
29	18	100.0	858	13	BQ436922	AGENCOURT
30	18	100.0	859	13	BU157686	AGENCOURT
31	18	100.0	861	9	AU133204	AU133204
32	18	100.0	861	13	BQ681189	AGENCOURT
33	18	100.0	868	12	BE762691	602734592
34	18	100.0	872	10	BE730286	601563427
35	18	100.0	872	13	BQ676711	AGENCOURT
36	18	100.0	872	13	BQ879010	AGENCOURT
37	18	100.0	877	9	AU131845	AU131845
38	18	100.0	877	13	BQ682391	AGENCOURT
39	18	100.0	877	13	BU557319	AGENCOURT
40	18	100.0	879	13	BU179560	AGENCOURT
41	18	100.0	886	13	BU160825	AGENCOURT
42	18	100.0	887	13	BU145169	AGENCOURT
43	18	100.0	888	10	BF026008	601670321
44	18	100.0	898	12	BI757143	603028814
45	18	100.0	899	10	BE277041	601178928

ALIGNMENTS

RESULT 1
AA376531

```
LOCUS       AA376531                 241 bp    mRNA    linear   EST 21-APR-1997
DEFINITION  EST88963 HSC172 cells II Homo sapiens cDNA 5' end similar to
            plasminogen activator, tissue type (GB:M15518), mRNA sequence.
ACCESSION   AA376531
VERSION     AA376531.1  GI:2028850
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 241)
AUTHORS     Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
```

TITLE   Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE   96026280
PUBMED   7566098
COMMENT   Other ESTs: THC167250
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.

FEATURES   Location/Qualifiers
     source   1..241
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="ATCC (inhost):180988"
                     /db_xref="taxon:9606"
                     /cell_type="fibroblast"
                     /cell_line="HSC172 (60PDL)"
                     /dev_stage="fetal"
                     /clone_lib="HSC172 cells II"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI"

BASE COUNT   54 a   70 c   68 g   46 t   3 others
ORIGIN

Query Match   100.0%; Score 18; DB 9; Length 241;
Best Local Similarity   100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGAC 18

RESULT 2
BE158218/c
LOCUS   BE158218   335 bp   mRNA   linear   EST 21-JUN-2000
DEFINITION MR2-HT0380-310100-102-g08 HT0380 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE158218
VERSION   BE158218.1 GI:8620939
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 335)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   MEDLINE 20202663
    PUBMED 10737800
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR2-HT0380-310
            100-102-g08&t3=2000-01-31&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 335.

FEATURES   Location/Qualifiers
     source   1..335
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0380"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

BASE COUNT   51 a   102 c   89 g   93 t
ORIGIN

Query Match   100.0%; Score 18; DB 10; Length 335;
Best Local Similarity   100.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
              ||||||||||||||||||
Db   31 TCTGAGGGAAACAGTGAC 14

RESULT 3
BE173847
LOCUS   BE173847   374 bp   mRNA   linear   EST 21-JUN-2000
DEFINITION CM0-HT0567-010300-256-g02 HT0567 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE173847
VERSION   BE173847.1 GI:8636573
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 374)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   MEDLINE 20202663
    PUBMED 10737800
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM0-HT0567-010
            300-256-g02&t3=2000-03-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 374.

FEATURES   Location/Qualifiers
     source   1..374

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HTO567"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT        68 a       116 c       123 g       67 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 10;   Length 374;
Matches    18;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    157 TCTGAGGGAAACAGTGAC 174

RESULT 4
BE173842
LOCUS       BE173842                378 bp    mRNA    linear    EST 21-JUN-2000
DEFINITION  CMO-HTO567-010300-256-e04 HTO567 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE173842
VERSION     BE173842.1  GI:8636568
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 378)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
PUBMED      10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2707001
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CMO-HTO567-010
            300-256-e04&t3=2000-03-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 378.
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HTO567"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
```

```
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT        68 a       116 c       126 g       68 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 10;   Length 378;
Matches    18;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    161 TCTGAGGGAAACAGTGAC 178

RESULT 5
BE259860
LOCUS       BE259860                476 bp    mRNA    linear    EST 26-OCT-2000
DEFINITION  601154266F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3510113 5',
            mRNA sequence.
ACCESSION   BE259860
VERSION     BE259860.1  GI:9130682
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 476)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM194 row: d column: 18
            High quality sequence stop: 475.
FEATURES             Location/Qualifiers
     source          1..476
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3510113"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_19"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Library constructed by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC library."
BASE COUNT       110 a       146 c       139 g       81 t
ORIGIN

Query Match
Best Local Similarity    100.0%;   Score 18;   DB 10;   Length 476;
Matches    18;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    182 TCTGAGGGAAACAGTGAC 199

RESULT 6
CB149125
LOCUS       CB149125                519 bp    mRNA    linear    EST 29-JAN-2003
DEFINITION  K-EST0205606 L15CKK1 Homo sapiens cDNA clone L15CKK1-35-D03 5',
```

```
            mRNA sequence.
ACCESSION   CB149125
VERSION     CB149125.1  GI:28131896
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
  COMMENT   Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 35  row: 03  column: 03
            High quality sequence stop: 519.
FEATURES             Location/Qualifiers
     source          1..519
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L15CKK1-35-D03"
                     /sex="M"
                     /cell_line="CK-K1"
                     /lab_host="Top10F'"
                     /clone_lib="L15CKK1"
                     /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tobacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."

BASE COUNT     122 a    141 c    165 g    91 t
ORIGIN

Query Match          100.0%;  Score 18;  DB 14;  Length 519;
Best Local Similarity 100.0%;  Pred. No. 4.5e+02;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    442 TCTGAGGGAAACAGTGAC 459

RESULT 7
BE386276
LOCUS       BE386276           582 bp    mRNA    linear    EST 21-JUL-2000
DEFINITION  601273495F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614782 5',
            mRNA sequence.
ACCESSION   BE386276
VERSION     BE386276.1  GI:9331641
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1  (bases 1 to 582)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM277 row: m column: 23
            High quality sequence stop: 582.
FEATURES             Location/Qualifiers
     source          1..582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3614782"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_20"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT     132 a    165 c    182 g    103 t
ORIGIN

Query Match          100.0%;  Score 18;  DB 10;  Length 582;
Best Local Similarity 100.0%;  Pred. No. 4.7e+02;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    470 TCTGAGGGAAACAGTGAC 487

RESULT 8
BE386111
LOCUS       BE386111           612 bp    mRNA    linear    EST 21-JUL-2000
DEFINITION  60127221F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3618149 5',
            mRNA sequence.
ACCESSION   BE386111
VERSION     BE386111.1  GI:9331476
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 612)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM286 row: j column: 06
            High quality sequence stop: 612.
FEATURES             Location/Qualifiers
     source          1..612
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
Db           439 TCTGAGGGAAACAGTGAC 456

RESULT 10
BG477932
LOCUS       BG477932                 644 bp    mRNA     linear    EST 21-MAR-2001
DEFINITION  602522911F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641386 5',
            mRNA sequence.
ACCESSION   BG477932
VERSION     BG477932.1  GI:13410211
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1408 row: e  column: 03
            High quality sequence stop: 639.
FEATURES             Location/Qualifiers
     source          1..644
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4641386"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_20"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT  151 a    183 c    196 g    112 t    2 others
ORIGIN
Query Match            100.0%;  Score 18;  DB 10;  Length 644;
Best Local Similarity  100.0%;  Pred. No. 4.9e+02;
Matches    18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
Db   490 TCTGAGGGAAACAGTGAC 507

RESULT 11
BE276291
LOCUS       BE276291                 682 bp    mRNA     linear    EST 13-JUL-2000
DEFINITION  601144209F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3050019 5',
            mRNA sequence.
ACCESSION   BE276291
VERSION     BE276291.1  GI:9151254
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3618149"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_20"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT  147 a    165 c    190 g    103 t    7 others
ORIGIN
Query Match            100.0%;  Score 18;  DB 10;  Length 612;
Best Local Similarity  100.0%;  Pred. No. 4.8e+02;
Matches    18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
Db   587 TCTGAGGGAAACAGTGAC 604

RESULT 9
AU117078
LOCUS       AU117078                 639 bp    mRNA     linear    EST 01-AUG-2002
DEFINITION  AU117078 HEMBA1 Homo sapiens cDNA clone HEMBA1000614 5', mRNA
            sequence.
ACCESSION   AU117078
VERSION     AU117078.1  GI:10932016
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 639)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished
  COMMENT   Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1..639
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEMBA1000614"
                     /clone_lib="HEMBA1"
                     /tissue_type="whole embryo, mainly head"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT  149 a    180 c    197 g    110 t    3 others
ORIGIN
Query Match            100.0%;  Score 18;  DB 9;  Length 639;
Best Local Similarity  100.0%;  Pred. No. 4.9e+02;
Matches    18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 TCTGAGGGAAACAGTGAC 18
         ||||||||||||||||||
```

```
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM95  row: b  column: 04
            High quality sequence stop: 614.
FEATURES             Location/Qualifiers
     source          1..682
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3050019"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_20"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT      151 a      194 c      216 g      121 t
ORIGIN

Query Match            100.0%;  Score 18;  DB 10;  Length 682;
Best Local Similarity  100.0%;  Pred. No. 5e+02;
Matches   18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    468 TCTGAGGGAAACAGTGAC 485

RESULT 12
BE895982
LOCUS       BE895982              690 bp    mRNA    linear  EST 20-OCT-2000
DEFINITION  601438641F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923602.5',
            mRNA sequence.
ACCESSION   BE895982
VERSION     BE895982.1  GI:11035926
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 690)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9760  row: a  column: 11
            High quality sequence stop: 582.
FEATURES             Location/Qualifiers
     source          1..690
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3923602"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2 kb.  Library constructed by Life
                     Technologies."

BASE COUNT      136 a      225 c      195 g      134 t
ORIGIN

Query Match            100.0%;  Score 18;  DB 10;  Length 690;
Best Local Similarity  100.0%;  Pred. No. 5e+02;
Matches   18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
Db    124 TCTGAGGGAAACAGTGAC 141

RESULT 13
BG476511
LOCUS       BG476511              711 bp    mRNA    linear  EST 21-MAR-2001
DEFINITION  6025220049F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4640172 5',
            mRNA sequence.
ACCESSION   BG476511
VERSION     BG476511.1  GI:13408790
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 711)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1405  row: b  column: 13
            High quality sequence stop: 711.
FEATURES             Location/Qualifiers
     source          1..711
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4640172"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_20"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT      173 a      189 c      225 g      123 t      1 others
ORIGIN

Query Match            100.0%;  Score 18;  DB 10;  Length 711;
Best Local Similarity  100.0%;  Pred. No. 5.1e+02;
Matches   18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TCTGAGGGAAACAGTGAC 18
          ||||||||||||||||||
```

Db   573 TCTGAGGGAAACAGTGAC 590

RESULT 14
BM042154
LOCUS   BM042154   725 bp   mRNA   linear   EST 07-NOV-2001
DEFINITION 6036157911P1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5520542 5',
           mRNA sequence.
ACCESSION   BM042154
VERSION   BM042154.1 GI:16771421
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 725)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: DCTD/DTP
           cDNA Library Preparation: Ling Hong/Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1875 row: e   column: 23
           High quality sequence stop: 725.
FEATURES   Location/Qualifiers
    source   1..725
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5520542"
              /tissue_type="melanotic melanoma, cell line"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_112"
              /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
              EcoRI; cDNA made by oligo-dT priming. Directionally cloned
              into EcoRI/XhoI sites using the following 5' adaptor:
              GGCACGAG(G). Library constructed by Ling Hong in the
              laboratory of Gerald M. Rubin (University of California,
              Berkeley) using ZAP-cDNA synthesis kit (Stracagene) and
              Superscript II RT (Life Technologies). Note: this is a
              NIH_MGC library."
BASE COUNT   178 a   203 c   219 g   125 t
ORIGIN

Query Match   100.0%;   Score 18; DB 12; Length 725;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
        |||||||||||||||||||
Db   502 TCTGAGGGAAACAGTGAC 519

RESULT 15
BI771956
LOCUS   BI771956   761 bp   mRNA   linear   EST 25-SEP-2001
DEFINITION 6030589811P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208117 5',
           mRNA sequence.
ACCESSION   BI771956
VERSION   BI771956.1 GI:15763534
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 761)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11522 row: b   column: 22
           High quality sequence stop: 738.
FEATURES   Location/Qualifiers
    source   1..761
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5208117"
              /lab_host="DH10B"
              /clone_lib="NIH_MGC_122"
              /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
              Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
              anonymous pool of 24 week female lung, 16 week female
              spleen, and 20-22 week male spleens. Library is oligo-dT
              primed and directionally cloned (EcoRV site is destroyed
              upon cloning). Average insert size 1.4 kb, insert size
              range 1-3 kb. Library is normalized and enriched for
              full-length clones and was constructed by C. Gruber
              (Invitrogen).   Research Genetics tracking code 026. Note:
              this is a NIH_MGC Library."
BASE COUNT   188 a   202 c   237 g   134 t
ORIGIN

Query Match   100.0%;   Score 18; DB 12; Length 761;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCTGAGGGAAACAGTGAC 18
        |||||||||||||||||||
Db   573 TCTGAGGGAAACAGTGAC 590

Search completed: August 19, 2003, 08:29:02
Job time : 59.6587 secs

THIS PAGE BLANK (USPTO)

d Aug 20 09:29:42 2003

OM nucleic - nucleic search, using sw model

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

Run on:         August 18, 2003, 19:53:48 ; Search time 277.124 Seconds
                                           (without alignments)
                                           9743.047 Million cell updates/sec

Title:          US-09-987-455-6
Perfect score:  66
Sequence:       1 atgaaaagacagctatcgc.........ctaccgtggcccaggcggcc 66

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters:       5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      GenEmbl:*
     1:  gb_ba:*
     2:  gb_htg:*
     3:  gb_in:*
     4:  gb_om:*
     5:  gb_ov:*
     6:  gb_pat:*
     7:  gb_ph:*
     8:  gb_pl:*
     9:  gb_pr:*
    10:  gb_ro:*
    11:  gb_sts:*
    12:  gb_sy:*
    13:  gb_un:*
    14:  gb_vi:*
    15:  em_ba:*
    16:  em_fun:*
    17:  em_hum:*
    18:  em_in:*
    19:  em_mu:*
    20:  em_om:*
    21:  em_or:*
    22:  em_ov:*
    23:  em_pat:*
    24:  em_ph:*
    25:  em_pl:*
    26:  em_ro:*
    27:  em_sts:*
    28:  em_un:*
    29:  em_vi:*
    30:  em_htg_hum:*
    31:  em_htg_inv:*
    32:  em_htg_other:*
    33:  em_htg_mus:*
    34:  em_htg_pln:*
    35:  em_htg_rod:*
    36:  em_htg_man:*
    37:  em_htg_vrt:*
    38:  em_sy:*
    39:  em_htgo_hum:*
    40:  em_htgo_mus:*
    41:  em_htgo_other:*
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	66	100.0	66	6	AX430829	AX430829 Sequence
2	66	100.0	66	6	AX431269	AX431269 Sequence
3	66	100.0	66	6	AX431272	AX431272 Sequence
4	66	100.0	705	12	HSIGKLC28	X95748 Synthetic c
5	66	100.0	705	12	HSIGKLC5	X95747 Synthetic c
6	66	100.0	711	12	HSIGKLC14	X95749 Synthetic c
7	66	100.0	711	12	HSIGKLG31	X95750 Synthetic c
8	66	100.0	1128	6	AX431268	AX431268 Sequence
9	66	100.0	1128	6	AX431271	AX431271 Sequence
10	66	100.0	3394	12	AF268280	AF268280 Phagemid
11	66	100.0	3450	6	AR241678	AR241678 Sequence
12	66	100.0	3758	12	AF268281	AF268281 Phagemid
13	66	100.0	5149	6	AX591634	AX591634 Sequence
14	66	100.0	5149	6	AX600065	AX600065 Sequence
15	66	100.0	5663	6	AX554424	AX554424 Sequence
16	66	100.0	6122	6	AX554413	AX554413 Sequence
17	66	100.0	6122	6	AX554422	AX554422 Sequence
18	66	100.0	6122	6	AX642149	AX642149 Sequence
19	66	100.0	6125	6	AX554420	AX554420 Sequence
20	64.4	97.6	5885	1	STAFRA	L42764 Staphylococc
21	61.2	92.7	580	6	AX742885	AX742885 Sequence
22	61.2	92.7	580	6	AX743509	AX743509 Sequence
23	60.2	91.2	74	6	E02814	E02814 DNA sequenc
24	60.2	91.2	74	6	E03955	E03955 DNA encodin
25	60.2	91.2	103	6	AX236721	AX236721 Sequence
26	60.2	91.2	249	6	I08200	I08200 Sequence 1
27	60.2	91.2	249	6	I52035	I52035 Sequence 1
28	60.2	91.2	438	12	SYNOMPARNA	M62710 Synthetic g
29	60.2	91.2	548	6	AR179433	AR179433 Sequence
30	60.2	91.2	548	6	AR282580	AR282580 Sequence
31	60.2	91.2	548	6	AX022441	AX022441 Sequence
32	60.2	91.2	548	6	AX030798	AX030798 Sequence
33	60.2	91.2	548	6	AX034619	AX034619 Sequence
34	60.2	91.2	652	6	E01685	E01685 hybrid DNA
35	60.2	91.2	739	1	AF234271	AF234271 Shigella
36	60.2	91.2	746	1	AF234270	AF234270 Salmonell
37	60.2	91.2	1400	1	STOMPA	X02006 Salmonella
38	60.2	91.2	1697	1	AF234269	AF234269 Escherich
39	60.2	91.2	2270	6	AX191724	AX191724 Sequence
40	60.2	91.2	2271	1	ECOMPA	V00307 E. coli omp
41	60.2	91.2	6477	6	A08458	A08458 pinf 4-49 m
42	60.2	91.2	7499	12	ASAK3121	AJ223121 Artificia
43	60.2	91.2	7500	12	ASAK3122	AJ223122 Artificia
c 44	60.2	91.2	11639	1	AE000198	AE000198 Escherich
c 45	60.2	91.2	11641	1	AE005286	AE005286 Escherich

ALIGNMENTS

```
RESULT 1
AX430829
LOCUS          AX430829               66 bp     DNA     linear    PAT  28-JUN-2002
DEFINITION     Sequence 1 from Patent WO0240696.
ACCESSION      AX430829
VERSION        AX430829.1   GI:21655908
KEYWORDS       .
SOURCE         Escherichia coli
  ORGANISM     Escherichia coli
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Escherichia.
REFERENCE      1
  AUTHORS      Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE        Methods for large scale protein production in prokaryotes
  JOURNAL      Patent: WO 0240696-A 1 23-MAY-2002;
```

Pred. No. is the number of results predicted by chance to have a

FEATURES   BOEHRINGER INGELHEIM INT (DE)
source   1..66
                  /organism="Escherichia coli"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:562"
BASE COUNT   14 a   18 c   21 g   13 t
ORIGIN

Query Match   100.0%;   Score 66;   DB 6;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 1.4e-11;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 2
AX431269
LOCUS   AX431269   66 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION   Sequence 3 from Patent WO0240650.
ACCESSION   AX431269
VERSION   AX431269.1   GI:21656151
KEYWORDS   .
SOURCE   Escherichia coli
  ORGANISM   Escherichia coli
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
               or k2s molecules
  JOURNAL   Patent: WO 0240650-A 3 23-MAY-2002;
               BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
source   1..66
                  /organism="Escherichia coli"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:562"
BASE COUNT   14 a   18 c   21 g   13 t
ORIGIN

Query Match   100.0%;   Score 66;   DB 6;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 1.4e-11;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 3
AX431272
LOCUS   AX431272   66 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION   Sequence 6 from Patent WO0240650.
ACCESSION   AX431272
VERSION   AX431272.1   GI:21656154
KEYWORDS   .
SOURCE   Escherichia coli
  ORGANISM   Escherichia coli
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.

  TITLE   Methods for large scale production of recombinant dna-derived tpa
               or k2s molecules
  JOURNAL   Patent: WO 0240650-A 6 23-MAY-2002;
               BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
source   1..66
                  /organism="Escherichia coli"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:562"
BASE COUNT   14 a   18 c   21 g   13 t
ORIGIN

Query Match   100.0%;   Score 66;   DB 6;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 1.4e-11;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 4
HSIGKLC28
LOCUS   HSIGKLC28   705 bp   mRNA   linear   SYN 29-MAR-2001
DEFINITION   Synthetic construct including Homo sapiens immunoglobulin kappa
               chain (clone: 28).
ACCESSION   X95748
VERSION   X95748.1   GI:1514578
KEYWORDS   constant region; immunoglobulin; kappa light chain.
SOURCE   synthetic construct
  ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1
  AUTHORS   Steinberger,P., Kraft,D. and Valenta,R.
  TITLE   Construction of a combinatorial IgE library from an allergic
               patient. Isolation and characterization of human IgE Fabs with
               specificity for the major timothy grass pollen allergen, Phl p 5
  JOURNAL   J. Biol. Chem. 271 (18), 10967-10972 (1996)
  MEDLINE   96210038
  PUBMED   8631916
REFERENCE   2 (bases 1 to 705)
  AUTHORS   Valenta,R.L.S.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
               Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
               1090 Vienna, AUSTRIA
FEATURES   Location/Qualifiers
source   1..705
                  /organism="synthetic construct"
                  /mol_type="mRNA"
                  /db_xref="taxon:32630"
                  /clone="28"
                  /cell_type="B-cell"
                  /clone_lib="pcomb3H"
                  1..705
                  /focus
source   1..66
                  /organism="Escherichia coli"
                  /mol_type="mRNA"
                  /db_xref="taxon:562"
source   67..705
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
gene   1..705
                  /gene="OMP-A"
CDS   1..705
                  /gene="OMP-A"
                  /codon_start=1
                  /transl_table=11

AUTHORS   Steinberger,P., Kraft,D., and Valenta,R.
TITLE   Construction of a combinatorial IgE library from an allergic
           patient. Isolation and characterization of human IgE Fabs with
           specificity for the major timothy grass pollen allergen, Phl p 5
JOURNAL   J. Biol. Chem. 271 (18), 10967-10972 (1996)
MEDLINE   96210038
PUBMED   8631916
REFERENCE 2 (bases 1 to 705)
AUTHORS   Valenta,R.L.S.
TITLE   Direct Submission
JOURNAL   Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
           Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
           1090 Vienna, AUSTRIA
FEATURES   Location/Qualifiers
     source   1..705
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /focus
     source   1..66
                     /organism="Escherichia coli"
                     /mol_type="mRNA"
                     /db_xref="taxon:562"
     source   67..705
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="5"
                     /cell_type="B-cell"
                     /clone_lib="pcomb3H"
     gene   1..705
                     /gene="OMP-A"
     CDS   1..705
                     /gene="OMP-A"
                     /codon_start=1
                     /transl_table=11
                     /product="immunoglobulin kappa light chain"
                     /protein_id="CAA65058.1"
                     /db_xref="GI:1514581"
                     /db_xref="REMTREMBL:CAA65058"
                     /translation="MKKTAIAIAVALAGFATVAQAAELTQSPSSLSASVGDRVTITCR
                     ASQGISSWLAWYQQKPGKAPKLLIYASSLQSGVPSRFSGSGSGTDFSLTISSLQPED
                     SATYYCQQANSFPYTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY
                     PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKLYACEVTH
                     QGLSSPVTKSFNRGEC"
     mRNA   1..66
                     /gene="OMP-A"
                     /note="leaderpeptide gene"
     misc_feature   67..129
                     /gene="OMP-A"
                     /note="FR1"
     primer_bind   67..84
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
     misc_feature   67..72
                     /gene="OMP-A"
                     /note="SAC 1 restriction site"
     misc_feature   130..162
                     /gene="OMP-A"
                     /note="CDR1"
     misc_feature   163..207
                     /gene="OMP-A"
                     /note="FR2"
     misc_feature   208..228
                     /gene="OMP-A"
                     /note="CDR2"
     misc_feature   229..324
                     /gene="OMP-A"
                     /note="FR3"
     misc_feature   325..351
                     /gene="OMP-A"
                     /note="CDR3"
     misc_feature   352..384

                     /product="immunoglobulin kappa light chain"
                     /protein_id="CAA65059.1"
                     /db_xref="GI:1514579"
                     /db_xref="REMTREMBL:CAA65059"
                     /translation="MKKTAIAIAVALAGFATVAQAAELTQSPSSLSASVGDRVTITCR
                     ASQSILGYLNWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTDFTLTITSLQPDD
                     FATYYCQQSYITPRTFGQGTKVEVKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY
                     PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTH
                     QGLSSPVTKSFNRGEC"
     mRNA   1..66
                     /gene="OMP-A"
                     /note="leaderpeptide gene"
     misc_feature   67..129
                     /gene="OMP-A"
                     /note="FR1"
     primer_bind   67..84
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
     misc_feature   67..72
                     /gene="OMP-A"
                     /note="CDR1"
     misc_feature   130..162
                     /gene="OMP-A"
                     /note="SAC 1 restriction site"
     misc_feature   163..207
                     /gene="OMP-A"
                     /note="FR2"
     misc_feature   208..228
                     /gene="OMP-A"
                     /note="CDR2"
     misc_feature   229..324
                     /gene="OMP-A"
                     /note="FR3"
     misc_feature   325..351
                     /gene="OMP-A"
                     /note="CDR3"
     misc_feature   352..384
                     /gene="OMP-A"
                     /note="FR4"
     C_region   385..702
                     /gene="OMP-A"
                     /note="kappa constant region"
     primer_bind   661..705
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
BASE COUNT   184 a   201 c   172 g   148 t
ORIGIN

Query Match   100.0%; Score 66; DB 12; Length 705;
Best Local Similarity   100.0%; Pred. No. 1.6e-11;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy 61 GCGGCC 66
        ||||||
Db 61 GCGGCC 66

RESULT 5
HSIGKLC5
LOCUS   HSIGKLC5   705 bp   mRNA   linear   SYN 29-MAR-2001
DEFINITION Synthetic construct including Homo sapiens immunoglobulin kappa
            chain (clone: 5).
ACCESSION   X95747
VERSION   X95747.1 GI:1514580
KEYWORDS   constant region; immunoglobulin; kappa light chain.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1

```
              /protein_id="CAA65060.1"
              /db_xref="GI:1514577"
              /db_xref="REMTREMBL;CAA65060"
              /translation="MKKTAIAIAVALAGFATVAQAAELVMTQSPSSLSASVGDRVTIT
              CRASHSISNYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQP
              EDFANYYCQSFSPSGTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNN
              FYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKLYACEV
              THQGLSSPVTKSFNRGEC"
     mRNA          1..66
              /gene="OMP-A"
              /note="leaderpeptide gene"
     misc_feature  67..135
              /gene="OMP-A"
              /note="FR1"
     primer_bind   67..84
              /gene="OMP-A"
              /note="PCR-primer sequence"
     misc_feature  67..72
              /gene="OMP-A"
              /note="SAC 1 restriction site"
     misc_feature  136..168
              /gene="OMP-A"
              /note="CDR1"
     misc_feature  169..213
              /gene="OMP-A"
              /note="FR2"
     misc_feature  214..234
              /gene="OMP-A"
              /note="CDR2"
     misc_feature  235..330
              /gene="OMP-A"
              /note="FR3"
     misc_feature  331..357
              /gene="OMP-A"
              /note="CDR3"
     misc_feature  358..390
              /gene="OMP-A"
              /note="FR4"
     C_region      391..708
              /gene="OMP-A"
              /note="kappa constant region"
     primer_bind   667..708
              /gene="OMP-A"
              /note="PCR-primer sequence"

BASE COUNT     187 a    194 c    179 g    151 t
ORIGIN

Query Match             100.0%;  Score 66;  DB 12;  Length 711;
Best Local Similarity   100.0%;  Pred. No. 1.6e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 7
HSIGKLG31
LOCUS       HSIGKLG31           711 bp    mRNA    linear   SYN 29-MAR-2001
DEFINITION  Synthetic construct including Homo sapiens immunoglobulin kappa
            chain (clone: 31).
ACCESSION   X95750
VERSION     X95750.1  GI:1514582
KEYWORDS    constant region; immunoglobulin; kappa light chain.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Steinberger,P., Kraft,D. and Valenta,R.
```

```
     C_region      /gene="OMP-A"
              /note="FR4"
              385..702
              /gene="OMP-A"
              /note="kappa constant region"
     primer_bind   661..705
              /gene="OMP-A"
              /note="PCR-primer sequence"

BASE COUNT     183 a    191 c    181 g    150 t
ORIGIN

Query Match             100.0%;  Score 66;  DB 12;  Length 705;
Best Local Similarity   100.0%;  Pred. No. 1.6e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 6
HSIGKLC14
LOCUS       HSIGKLC14           711 bp    mRNA    linear   SYN 29-MAR-2001
DEFINITION  Synthetic construct including Homo sapiens immunoglobulin kappa
            chain (clone: 14).
ACCESSION   X95749
VERSION     X95749.1  GI:1514576
KEYWORDS    constant region; immunoglobulin; kappa light chain.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Steinberger,P., Kraft,D. and Valenta,R.
  TITLE     Construction of a combinatorial IgE library from an allergic
            patient. Isolation and characterization of human IgE Fabs with
            specificity for the major timothy grass pollen allergen, Phl p 5
  JOURNAL   J. Biol. Chem. 271 (18), 10967-10972 (1996)
  MEDLINE   96210038
   PUBMED   8631916
REFERENCE   2  (bases 1 to 711)
  AUTHORS   Valenta,R.L.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
            Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
            1090 Vienna, AUSTRIA
FEATURES             Location/Qualifiers
     source        1..711
              /organism="synthetic construct"
              /mol_type="mRNA"
              /db_xref="taxon:32630"
              /focus
     source        1..66
              /organism="Escherichia coli"
              /mol_type="mRNA"
              /db_xref="taxon:562"
     source        67..711
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="14"
              /cell_type="B-cell"
              /clone_lib="pcomb3H"
     gene          1..711
              /gene="OMP-A"
     CDS           1..711
              /gene="OMP-A"
              /codon_start=1
              /transl_table=11
              /product="immunoglobulin kappa light chain"
```

TITLE   Construction of a combinatorial IgE library from an allergic
            patient. Isolation and characterization of human IgE Fabs with
            specificity for the major timothy grass pollen allergen, Phl p 5
JOURNAL   J. Biol. Chem. 271 (18), 10967-10972 (1996)
MEDLINE   96210038
PUBMED   8631916
REFERENCE   2 (bases 1 to 711)
AUTHORS   Valenta,R.L.S.
TITLE   Direct Submission
JOURNAL   Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
            Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
            1090 Vienna, AUSTRIA
FEATURES   Location/Qualifiers
     source   1..711
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /focus
     source   1..66
                     /organism="Escherichia coli"
                     /mol_type="mRNA"
                     /db_xref="taxon:562"
     source   67..711
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="31"
                     /cell_type="B-cell"
                     /clone_lib="pcomb3H"
     gene   1..711
                     /gene="OMP-A"
     CDS   1..711
                     /gene="OMP-A"
                     /codon_start=1
                     /transl_table=11
                     /product="immunoglobulin kappa light chain"
                     /protein_id="CAA65061.1"
                     /db_xref="GI:1514583"
                     /db_xref="REMTREMBL:CAA65061"
                     /translation="MKKTAIAIAVALAGFATVAQAAELVMTQSPSSLSAFVGDRVTIT
                     CRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGVPSRFSGSGSGTDFTLTISSLQP
                     EDFATYYCQQSHSTPYTFGGGTNLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNN
                     FYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEV
                     THQGLSSPVTKSFNRGEC"
     mRNA   1..66
                     /gene="OMP-A"
                     /note="leaderpeptide gene"
     misc_feature   67..135
                     /gene="OMP-A"
                     /note="FR1"
     primer_bind   67..84
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
     misc_feature   67..72
                     /gene="OMP-A"
                     /note="SAC 1 restriction site"
     misc_feature   136..168
                     /gene="OMP-A"
                     /note="CDR1"
     misc_feature   169..213
                     /gene="OMP-A"
                     /note="FR2"
     misc_feature   214..234
                     /gene="OMP-A"
                     /note="CDR2"
     misc_feature   235..330
                     /gene="OMP-A"
                     /note="FR3"
     misc_feature   331..357
                     /gene="OMP-A"
                     /note="CDR3"
     misc_feature   358..390
                     /gene="OMP-A"

     C_region   /note="FR4"
                     391..708
                     /gene="OMP-A"
                     /note="kappa constant region"
     primer_bind   667..708
                     /gene="OMP-A"
                     /note="PCR-primer sequence"

BASE COUNT   188 a   194 c   179 g   150 t
ORIGIN

Query Match   100.0%; Score 66; DB 12; Length 711;
Best Local Similarity   100.0%; Pred. No. 1.6e-11;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66

RESULT 8
AX431268
LOCUS   AX431268   1128 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 2 from Patent WO0240650.
ACCESSION   AX431268
VERSION   AX431268.1 GI:21656150
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 2 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
     source   1..1128
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for OmpA-K2S fusion protein"

BASE COUNT   238 a   332 c   332 g   226 t
ORIGIN

Query Match   100.0%; Score 66; DB 6; Length 1128;
Best Local Similarity   100.0%; Pred. No. 1.6e-11;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66

RESULT 9
AX431271
LOCUS   AX431271   1128 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 5 from Patent WO0240650.
ACCESSION   AX431271
VERSION   AX431271.1 GI:21656153
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.

```
TITLE     Methods for large scale production of recombinant dna-derived tpa
          or k2s molecules
JOURNAL   Patent: WO 0240650-A 5 23-MAY-2002;
          BOEHRINGER INGELHEIM INT (DE)
FEATURES
  source          1..1128
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
                  /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT      238 a   332 c   332 g   226 t
ORIGIN

Query Match     100.0%;  Score 66;  DB 6;  Length 1128;
Best Local Similarity   100.0%;  Pred. No. 1.6e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 10
AF268280
LOCUS       AF268280              3394 bp    DNA    circular SYN 03-OCT-2000
DEFINITION  Phagemid cloning vector pComb3H, complete sequence.
ACCESSION   AF268280
VERSION     AF268280.1  GI:10505048
KEYWORDS    .
SOURCE      Phagemid cloning vector pComb3H
  ORGANISM  Phagemid cloning vector pComb3H
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 3394)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  JOURNAL   (in) PHAGE DISPLAY, A LABORATORY MANUAL, Cold Spring Harbor
            Laboratory Press, Cold Spring Harbor, NY, USA (2000). In press
REFERENCE   2  (bases 1 to 3394)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,
            Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA
            92037, USA
FEATURES
  source          1..3394
                  /organism="Phagemid cloning vector pComb3H"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:137785"
BASE COUNT      862 a   783 c   876 g   873 t
ORIGIN

Query Match     100.0%;  Score 66;  DB 12;  Length 3394;
Best Local Similarity   100.0%;  Pred. No. 1.7e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 281

Qy   61 GCGGCC 66
        ||||||
Db  282 GCGGCC 287

RESULT 11
AR241678
LOCUS       AR241678              3450 bp    DNA    linear   PAT 20-DEC-2002
DEFINITION  Sequence 20 from patent US 6472147.
ACCESSION   AR241678
VERSION     AR241678.1  GI:27287480
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3450)
  AUTHORS   Janda,K.D., Wirsching,P., Lerner,R.A. and Gao,C.
  TITLE     Methods for display of heterodimeric proteins on filamentous phage
            using pVII and pIX, compositions, vectors and combinatorial
            libraries
  JOURNAL   Patent: US 6472147-A 20 29-OCT-2002;
FEATURES                                         Location/Qualifiers
  source          1..3450
                  /organism="unknown"
BASE COUNT      873 a   799 c   891 g   887 t
ORIGIN

Query Match     100.0%;  Score 66;  DB 6;  Length 3450;
Best Local Similarity   100.0%;  Pred. No. 1.7e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2667 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2726

Qy    61 GCGGCC 66
         ||||||
Db  2727 GCGGCC 2732

RESULT 12
AF268281
LOCUS       AF268281              3758 bp    DNA    circular SYN 03-OCT-2000
DEFINITION  Phagemid cloning vector pComb3X, complete sequence.
ACCESSION   AF268281
VERSION     AF268281.1  GI:10505049
KEYWORDS    .
SOURCE      Phagemid cloning vector pComb3X
  ORGANISM  Phagemid cloning vector pComb3X
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 3758)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  JOURNAL   (in) PHAGE DISPLAY, A LABORATORY MANUAL, Cold Spring Harbor
            Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press
REFERENCE   2  (bases 1 to 3758)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,
            Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA
            92037, USA
FEATURES                                         Location/Qualifiers
  source          1..3758
                  /organism="Phagemid cloning vector pComb3X"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:137786"
BASE COUNT      939 a   896 c   959 g   964 t
ORIGIN

Query Match     100.0%;  Score 66;  DB 12;  Length 3758;
Best Local Similarity   100.0%;  Pred. No. 1.7e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 281

Qy   61 GCGGCC 66
        ||||||
Db  282 GCGGCC 287

RESULT 13
AX591634
LOCUS       AX591634              5149 bp    DNA    linear   PAT 27-JAN-2003
```

```
DEFINITION   Sequence 8 from Patent WO0246434.
ACCESSION    AX591634
VERSION      AX591634.1  GI:27950030
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    Bowdish,K.S. and Barbas-Frederickson,S.
  TITLE      Novel plasmid vectors
  JOURNAL    Patent: WO 0246434-A 8 13-JUN-2002;
             ALEXION PHARMACEUTICALS, INC. (US)
FEATURES          Location/Qualifiers
  source          1..5149
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
                  /note="vector"
BASE COUNT     1279 a   1244 c   1372 g   1254 t
ORIGIN

Query Match       100.0%;  Score 66; DB 6; Length 5149;
Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
Matches    66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676
```

```
RESULT 14
AX600065
LOCUS        AX600065           5149 bp    DNA        linear    PAT 14-FEB-2003
DEFINITION   Sequence 60 from Patent WO0246238.
ACCESSION    AX600065
VERSION      AX600065.1  GI:28400138
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    Bowdish,K.S., Barbas-Frederickson,S. and Renshaw,M.
  TITLE      Rationally designed antibodies
  JOURNAL    Patent: WO 0246238-A 60 13-JUN-2002;
             ALEXION PHARMACEUTICALS, INC. (US)
FEATURES          Location/Qualifiers
  source          1..5149
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
                  /note="vector"
BASE COUNT     1279 a   1244 c   1372 g   1254 t
ORIGIN

Query Match       100.0%;  Score 66; DB 6; Length 5149;
Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
Matches    66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676
```

```
RESULT 15
AX554424
LOCUS        AX554424           5683 bp    DNA        linear    PAT 27-NOV-2002
DEFINITION   Sequence 12 from Patent WO0246436.
ACCESSION    AX554424
VERSION      AX554424.1  GI:25898200
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    Bowdish,K.S., Barbas-Frederickson,S., Wild,M. and Mcwhirter,J.
  TITLE      Novel plasmid vectors
  JOURNAL    Patent: WO 0246436-A 12 13-JUN-2002;
             ALEXION PHARMACEUTICALS, INC. (US)
FEATURES          Location/Qualifiers
  source          1..5683
                  /organism="synthetic construct"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:32630"
                  /note="plasmid"
BASE COUNT     1337 a   1492 c   1543 g   1311 t
ORIGIN

Query Match       100.0%;  Score 66; DB 6; Length 5683;
Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
Matches    66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676
```

```
Search completed: August 19, 2003, 04:46:35
Job time : 278.124 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      August 18, 2003, 19:51:53 ; Search time 23.2513 seconds
                                    (without alignments)
                         7662.485 Million cell updates/sec

Title:        US-09-987-455-6
Perfect score: 66
Sequence:      1 atgaaaaagacagctatcgc.........ctaccgtggccaggcggcc 66

Scoring table: IDENTITY NUC
              Gapop 10.0 , Gapext 1.0

Searched:      252756 seqs, 134971901 residues

Total number of hits satisfying chosen parameters:      5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_19Jun03:*
      1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
      2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
      3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
      4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
      5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
      6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
      7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
      8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
      9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
     10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
     11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
     12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
     13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
     14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
     15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
     16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
     17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
     18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
     19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
     20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
     21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
     22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
     23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
     24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
     25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result      %    Query
No.  Score Match Length DB    ID         Description
----------------------------------------------------------------------
1    66   100.0  66    24  AAD40615   Escherichia coli o
2    66   100.0  66    24  AAD38987   Escherichia coli O
3    66   100.0  1128  24  AAD40614   OmpA-K2S fusion pr
4    66   100.0  1574  21  AAZ57599   Erythrocyte and He
5    66   100.0  3450  22  AAC66993   Vector pCGMT-1b.
6    66   100.0  4691  16  AAQ92546   pComb3 expression
7    66   100.0  5149  24  ABQ73372   Plasmid pRL8 nucle
8    66   100.0  5149  24  ABN84077   Phagemid vector pR
9    66   100.0  5683  24  AAL41116   Plasmid pRL5 nucle
10   66   100.0  6122  24  AAL41112   Plasmid pRL5 CAT n
11   66   100.0  6122  24  AAL41115   Plasmid pRL5 bsi-C
12   66   100.0  6122  24  ABN84078   Plasmid pRL5-CAT.
13   66   100.0  6125  24  AAL41114   Plasmid pRL5 CAT-A
14   66   100.0  6166  16  AAQ92547   Expression vector,
15   60.2  91.2  73    12  AAQ10569   Omp A signal pepti
16   60.2  91.2  103   22  AAS44154   Neisseria meningit
17   60.2  91.2  249   11  AAQ03521   TAC recombinant DN
18   60.2  91.2  548   18  AAT91586   Growth factor LHL
19   60.2  91.2  548   20  AAX34590   DNA encoding a LHL
20   60.2  91.2  652    9  AAN81344   Sequence encoding
21   60.2  91.2  2814  14  AAQ41731   PE binding/translo
22   60.2  91.2  2814  14  AAQ38414   PE binding and tra
23   60.2  91.2  6477  11  AAQ02030   Plasmid pinf 4-49.
24   59.6  90.3  73    21  AAA12579   Synthetic DNA enco
25   59.6  90.3  415   17  AAT15982   pSKAN8 fragment #1
26   59.6  90.3  470   18  AAT91590   Kappa light chain
27   59.6  90.3  470   20  AAX34594   DNA encoding a hum
28   59.6  90.3  599   18  AAT91588   Growth factor LHL.
29   59.6  90.3  599   20  AAX34593   DNA encoding a for
30   59.6  90.3  759   19  AAV53721   The ORF of the clo
31   59.6  90.3  770   19  AAV53720   Cloned duplicate o
32   59.6  90.3  793   20  AAX24101   Plasmid pBBP22 DNA
33   59.6  90.3  793   22  AAF25703   Bilin binding-prot
34   59.6  90.3  803   12  AAQ14831   Qm212 single chain
35   59.6  90.3  932   22  AAA91151   Expression cassett
36   59.6  90.3  1031  18  AAT91589   Growth factor TLHL
37   59.6  90.3  1031  20  AAX34592   DNA encoding TLHL.
38   59.6  90.3  1219  20  AAX24099   Plasmid pBBP20 DNA
39   59.6  90.3  1219  22  AAF25695   Bilin binding-prot
40   59.6  90.3  1219  20  AAF25708   Bilin binding-prot
41   59.6  90.3  1380  20  AAX24100   Plasmid pBBP21 DNA
42   59.6  90.3  1380  22  AAF25710   Bilin binding-prot
43   59.6  90.3  1479  20  AAX34596   DNA encoding a ccM
44   59.6  90.3  1490  18  AAT91587   Growth factor CATA
45   59.6  90.3  1490  20  AAX34591   DNA encoding CATAB
```

ALIGNMENTS

```
RESULT 1
AAD40615
ID   AAD40615 standard; DNA; 66 BP.
XX
AC   AAD40615;
XX
DT   30-OCT-2002  (first entry)
XX
DE   Escherichia coli ompA peptide encoding DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..66
FT                   /*tag= a
FT                   /product= "OmpA peptide"
FT                   /note= "No stop codon"
FT                   /partial
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
```

XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25035.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 8; Page 30; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA fragment encoding Escherichia coli OmpA peptide.
XX
SQ   Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;

Query Match   100.0%;   Score 66;   DB 24;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 2.1e-14;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
         ||||||
Db   61 GCGGCC 66

RESULT 2
AAD38987
ID   AAD38987 standard; DNA; 66 BP.
XX
AC   AAD38987;
XX
DT   23-SEP-2002 (first entry)
XX
DE   Escherichia coli OmpA DNA.
XX
KW   Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;
KW   serine protease; ds.
XX
OS   Escherichia coli.
XX
PN   WO20020696-A2.
XX
PD   23-MAY-2002.
XX
PF   08-NOV-2001; 2001WO-EP12920.
XX
PR   14-NOV-2000; 2000GB-0027782.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-471625/50.
XX

PT   Producing recombinant DNA-derived kringle 2 plus serine protease,
PT   comprises using a prokaryotic cell expressing a vector having a DNA
PT   coding for a heterologous protein operably linked to a DNA coding for
PT   the signal peptide OmpA -
XX
PS   Claim 9; Page 23; 52pp; English.
XX
CC   The invention relates to a method for producing recombinant DNA-derived
CC   heterologous protein in prokaryotic cells, where the heterologous protein
CC   is secreted extracellularly in an active and correctly folded protein and
CC   the prokaryotic cell contains and expresses a vector comprising the DNA
CC   coding for the heterologous protein operably linked to the DNA coding
CC   for the signal peptide OmpA or its functional derivative. The method is
CC   useful for commercial large-scale production of heterologous proteins,
CC   e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is
CC   generally applicable in the expression of several different proteins and
CC   polypeptides which do not require mammalian glycosylation in prokaryotic
CC   host cells. The method may also be used to obtain DNA sequences of a
CC   protein of interest to be expressed from databases and cloned for use.
CC   The present sequence is Escherichia coli OmpA DNA.
XX
SQ   Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;

Query Match   100.0%;   Score 66;   DB 24;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 2.1e-14;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
         ||||||
Db   61 GCGGCC 66

RESULT 3
AAD40614
ID   AAD40614 standard; DNA; 1128 BP.
XX
AC   AAD40614;
XX
DT   30-OCT-2002 (first entry)
XX
DE   OmpA-K2S fusion protein encoding DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; fusion protein; gene; ds.
XX
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1128
FT   /*tag= a
FT   /product= "OmpA-K2S fusion protein"
FT   /transl_except= (pos:1126..1128, aa:Pro-Gly)
FT   /note= "No stop codon"
FT   /partial
XX
PN   WO20020650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi C, Manosroi J, Manosroi A;
XX

DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25034.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 7; Page 30; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA encoding a fusion protein comprising OmpA and K2S protein.
XX
SQ   Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;

Query Match   100.0%; Score 66; DB 24; Length 1128;
Best Local Similarity 100.0%; Pred. No. 4.3e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGGTACCGTGGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGGTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
              ||||||
Db   61 GCGGCC 66

RESULT 4
AAZ57599
ID   AAZ57599 standard; DNA; 1574 BP.
XX
AC   AAZ57599;
XX
DT   19-APR-2000 (first entry)
XX
DE   Erythrocyte and Hepatitis B dual-specific antibody DNA sequence.
XX
KW   Erythrocyte; Hepatitis B; dual-specific antibody; surface antigen; ds.
XX
OS   Unidentified.
XX
PN   CN1232039-A.
XX
PD   20-OCT-1999.
XX
PF   02-APR-1999; 99CN-0103517.
XX
PR   02-APR-1999; 99CN-0103517.
XX
PA   (NAVA-) NAVAL GEN HOSPITAL PLA.
XX
PI   Chen Y, Wang Y;
XX
DR   WPI; 2000-098467/09.
XX
PT   Genetic engineering double specific antibody and its use -
XX
PS   Claim 1-2; 6pp; Chinese.
XX
CC   The present invention describes a dual-specific antibody composed of
CC   the surface antigens to resist against erythrocyte and hepatitis B,

CC   which can be used as the test reagent to detect the surface antigen
CC   of hepatitis B in blood. The dual-specific antibody is prepared by
CC   recombination technique in gene engineering. The recombination technique
CC   includes reforming the surface antigen to resist against erythrocyte
CC   and hepatitis B to become hybrid antibody genes by shortening the
CC   joining peptide of single-chain antibody, and assembling them in the
CC   same expression vector. The gene product can be directly extracted from
CC   the supernatant of bacterial culture liquid. Its advantages are low
CC   cost, quick detection and simple operation.
XX
SQ   Sequence 1574 BP; 395 A; 390 C; 419 G; 370 T; 0 other;

Query Match   100.0%; Score 66; DB 21; Length 1574;
Best Local Similarity 100.0%; Pred. No. 4.6e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGGTACCGTGGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   792 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGGTACCGTGGCCCAG 851

Qy   61 GCGGCC 66
              ||||||
Db   852 GCGGCC 857

RESULT 5
AAC66993
ID   AAC66993 standard; DNA; 3450 BP.
XX
AC   AAC66993;
XX
DT   28-MAR-2001 (first entry)
XX
DE   Vector pCGMT-1b.
XX
KW   Filamentous phage; protein display; pVII; pIX;
KW   combinatorial antibody library; ds.
XX
OS   Synthetic.
XX
PN   WO200071694-A1.
XX
PD   30-NOV-2000.
XX
PF   24-MAY-2000; 2000WO-US14433.
XX
PR   25-MAY-1999; 99US-0318786.
XX
PA   (SCRI) SCRIPPS RES INST.
XX
PI   Janda KD, Wirsching P, Lerner RA, Gao C;
XX
DR   WPI; 2001-032030/04.
XX
PT   Novel filamentous phage encapsulating a genome encoding fusion
PT   polypeptide comprising exogenous polypeptide fused to amino terminus of
PT   pVII and pIX proteins, for constructing diverse heterodimeric
PT   polypeptide array -
XX
PS   Example 1; Page 74-75; 90pp; English.
XX
CC   The present invention describes a filamentous phage encapsulating a
CC   genome encoding a fusion protein. This fusion protein comprises an
CC   exogenous protein fused to the amino terminus of a filamentous phage pVII
CC   or pIX protein. This is useful in the design of proteins for medical,
CC   industrial, environmental and research applications.
XX
SQ   Sequence 3450 BP; 873 A; 799 C; 891 G; 887 T; 0 other;

Query Match   100.0%; Score 66; DB 22; Length 3450;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2667 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2726

Qy   61 GCGGCC 66
          ||||||
Db   2727 GCGGCC 2732

RESULT 6
AAQ92546
ID   AAQ92546 standard; DNA; 4691 BP.
XX
AC   AAQ92546;
XX
DT   11-MAR-1996 (first entry)
XX
DE   pComb3 expression vector.
XX
KW   pComb3; phagemid expression vector; bacteriophage; coat protein 3;
KW   Gene III; filamentous phage; minor phage coat protein; cpIII; cp3;
KW   bacterial membrane; periplasm; E. coli; human; Fab; HIV; gp120;
KW   combinatorial Fab library; cassette; Fd/cp3; lacZ promoter/operator;
KW   ribosome binding site; RBS; PelB leader; spacer; tether sequence;
KW   MT4; pMT4-3; antibody; ss; cyclic.
XX
OS   Synthetic.
XX
PN   WO9511317-A1.
XX
PD   27-APR-1995.
XX
PF   19-OCT-1994; 94WO-US11907.
XX
PR   19-SEP-1994; 94US-0308841.
PR   19-OCT-1993; 93US-0139409.
PR   26-APR-1994; 94US-0233619.
XX
PA   (SCRI) SCRIPPS RES INST.
XX
PI   Barbas CF, Burton DR, Lerner RA;
XX
DR   WPI; 1995-170235/22.
XX
PT   Synthetic human neutralising monoclonal antibodies to human
PT   immunodeficiency virus - used for diagnosis and immuno:therapy of
PT   HIV-induced disease
XX
PS   Example 1; Page 185-188; 249pp; English.
XX
CC   This sequence represents the pComb3 phagemid expression vector. This
CC   vector has been designed to allow for anchoring of expressed proteins on
CC   the bacteriophage coat protein 3. Gene III of filamentous phage encodes
CC   the 406 residue minor phage coat protein, cpIII (cp3), which is
CC   expressed prior to extrusion in the phage assembly process on a bacterial
CC   membrane and accumulates on the inner membrane facing into the periplasm
CC   of E. coli. This plasmid was used within the scope of the invention to
CC   express various mutagenised human Fab's which comprise heavy and light
CC   variable regions which bind to HIV gp120. pComb3 allows for both surface
CC   display and soluble forms of the Fabs. The vector was designed for the
CC   cloning of combinatorial Fab libraries. pComb consists of a DNA molecule
CC   having two cassettes to express one fusion protein, Fd/cp3, and one
CC   soluble protein, the light chain. The finished vector comprises,
CC   operatively linked 5' to 3', a first cassette consisting of lacZ
CC   promoter/operator sequences, a NotI restriction site, a ribosome binding
CC   site (RBS), a PelB leader, a spacer region, a cloning region bordered by
CC   5' XhoI and 3' SpeI restriction sites, the tether sequence, the sequences
CC   encoding bacteriophage cp3 followed by a stop codon, a NheI restriction
CC   site between the two cassettes, and a second lacZ promoter/operator
CC   sequence, followed by an expression control RBS, a PelB leader, a spacer
CC   region, a cloning region bordered by 5' SacI and 3' XbaI restriction
CC   sites, followed by expression control stop sequences and a second NotI
CC   restriction site. The pComb3 expression vector forms the basic

CC   construct of the MT4 Fab display phagemid expression vector, pMT4-3 (see
CC   also AAQ92540), used in the invention for the production of synthetic
CC   human Fab antibodies against gp120 of HIV.
XX
SQ   Sequence 4691 BP; 1170 A; 1171 C; 1232 G; 1118 T; 0 other;

Query Match   100.0%; Score 66; DB 16; Length 4691;
Best Local Similarity   100.0%; Pred. No. 6.1e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676

RESULT 7
ABQ73372
ID   ABQ73372 standard; DNA; 5149 BP.
XX
AC   ABQ73372;
XX
DT   01-OCT-2002 (first entry)
XX
DE   Plasmid pRL8 nucleotide sequence SEQ ID NO:60.
XX
KW   TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW   complementarity determining region; immunoglobin; antianaemic;
KW   haemostatic; nephrotropic; haematopoietic cell; haematopoiesis;
KW   gene; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO20024623B-A2.
XX
PD   13-JUN-2002.
XX
PF   05-DEC-2001; 2001WO-US47656.
XX
PR   05-DEC-2000; 2000US-251448P.
PR   04-MAY-2001; 2001US-288889P.
PR   29-MAY-2001; 2001US-294068P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-Frederickson S, Renshaw M;
XX
DR   WPI; 2002-566610/60.
XX
PT   A novel immunogen molecule comprising a region in which amino acid
PT   residues corresponding to at least a portion of the complementary
PT   determining region are replaced or fused with an erythropoietin or
PT   thrombopoetin mimetic -
XX
PS   Disclosure; Fig 6A-C; 113pp; English.
XX
CC   The present invention describes an immunoglobin molecule or its fragment
CC   (I) comprising a region where amino acid residues corresponding to at
CC   least a portion of the complementary determining region (CDR) are
CC   replaced or fused with biologically active peptides e.g. a peptide
CC   mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC   that is flanked with proline at its carboxy terminus. (I) has
CC   antianaemic, haemostatic and nephrotropic activities, and can be used as
CC   a stimulator of proliferation, differentiation and maturation of
CC   haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC   for stimulating proliferation, differentiation or growth of
CC   promegakaryocytes or megakaryocytes, where (I) is contacted with
CC   promegakaryocytes or megakaryocytes, which results in increased platelet
CC   production. (I) with a region where amino acid residues corresponding to

CC   a portion of CDR is replaced with an EPO mimetic, or which has one or
CC   more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC   production of red blood cells, where (I) is contacted with haematopoietic
CC   stem cells or their progenitors. (I) is useful for diagnostics or
CC   therapeutics, in cell isolation strategies, and for treating patients
CC   suffering from deficiency in cell populations caused by disease,
CC   disorders or treatments related to the suppression of haematopoiesis.
CC   ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
CC   in the exemplification of the present invention.
XX
SQ   Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match   100.0%; Score 66; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 6.2e-14;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2611 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
         ||||||
Db 2671 GCGGCC 2676

RESULT 8
ABN84077
ID   ABN84077 standard; DNA; 5149 BP.
XX
AC   ABN84077;
XX
DT   23-SEP-2002 (first entry)
XX
DE   Phagemid vector pRL8.
XX
KW   Vector; phagemid; pRL8; pComb; dimerisation domain; Jun;
KW   leucine zipper; antibody; gene; ds.
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   2677..3866
FT   /*tag= a
FT   /note= "light chain stuffer region"
FT   misc_feature   3968..4337
FT   /*tag= b
FT   /note= "heavy chain stuffer region"
FT   CDS   4338..4583
FT   /*tag= c
FT   /partial
FT   /product= "Jun dimerisation domain"
FT   /note= "the CDS does not include a start codon"
XX
PN   WO20246434-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46514.
XX
PR   08-DEC-2000; 2000US-254410P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-Frederickson S;
XX
DR   WPI; 2002-537568/57.
DR   P-PSDB; ABB79463.
XX
PT   Novel pComb family phagemid comprising a nucleic acid encoding a
PT   dimerization domain, useful for producing antibody fragments which
PT   require dimerization in order to crosslink for activation of target
PT   receptors -
XX

PS   Disclosure; Fig 8A-C; 17pp; English.
XX
CC   The present sequence is of plasmid pLR8, a pComb vector family
CC   member. pLR8 was constructed by insertion of a PCR fragment
CC   containing a flexible linker (murine kappa hinge region) followed
CC   by a Jun leucine zipper dimerisation domain into a pRL4 backbone
CC   that had been treated to remove the stuffer region. The resulting
CC   construct was then cut with restriction endonucleases to replace
CC   scFv with a pRL4 stuffer region for the light and heavy chain. The
CC   vector is useful in the production of biologically active proteins
CC   such as antibody fragments which require dimerisation for their
CC   biological activity. It can also be used in the production and
CC   screening of libraries made in accordance with phage display
CC   technology. Homo-dimerisation of single chain antibodies is
CC   accomplished by the presence of the Jun leucine zipper domain,
CC   which is responsible for protein-protein interactions. Subcloning
CC   to pRL8 individually or en masse following FACS sorting or panning
CC   allows expression, e.g., of dimeric soluble Fabs.
XX
SQ   Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match   100.0%; Score 66; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 6.2e-14;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2611 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
         ||||||
Db 2671 GCGGCC 2676

RESULT 9
AAL41116
ID   AAL41116 standard; DNA; 5683 BP.
XX
AC   AAL41116;
XX
DT   16-OCT-2002 (first entry)
XX
DE   Plasmid pRL5 nucleic acid sequence.
XX
KW   Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW   cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW   pRL5-CAT; pRL5 bsi-CAT; gene; ds.
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   591..1451
FT   /*tag= a
FT   /product= "Beta lactamase"
FT   CDS   2611..2676
FT   /*tag= b
FT   CDS   3508..3828
FT   /*tag= c
FT   /product= "Omp A leader amino acid sequence"
FT   CDS   3859..3924
FT   /*tag= d
FT   /product= "Kappa constant region amino acid sequence"
FT   CDS   4740..5651
FT   /*tag= e
FT   /product= "pel B leader amino acid sequence"
FT   /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT   sequence"
XX
PN   WO20246436-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46516.

XX
PR   08-DEC-2000; 2000US-254411P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;
XX
DR   WPI; 2002-537570/57.
DR   P-PSDB; AAO22536, AAO22537, AAO22538, AAO22539, AAO22542.
XX
PT   Novel plasmid useful in cloning and expression of foreign genetic
PT   information -
XX
PS   Claim 1; Fig 4A-E; 39pp; English.
XX
CC   The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC   6125, or 5683 base pairs fully defined in the specification. The
CC   invention more specifically relates to novel vectors capable of
CC   replication and expression of foreign genetic information in bacteria,
CC   such as, for example, cyanobacterium and E. coli. The new vectors have
CC   been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC   new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC   polynucleotide sequence represents the plasmid pRL5 nucleic acid sequence
CC   of the invention.
XX
SQ   Sequence 5683 BP; 1337 A; 1495 C; 1540 G; 1311 T; 0 other;

```
Query Match             100.0%;  Score 66;  DB 24;  Length 5683;
Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy         61 GCGGCC 66
              ||||||
Db       2671 GCGGCC 2676
```

RESULT 10
AAL41112
ID   AAL41112 standard; DNA; 6122 BP.
XX
AC   AAL41112;
XX
DT   16-OCT-2002 (first entry)
XX
DE   Plasmid pRL5 CAT nucleic acid sequence.
XX
KW   Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW   cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW   pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1013..1672
FT   /tag= a
FT   /product= "CAT amino acid sequence"
FT   CDS   3050..3115
FT   /tag= b
FT   CDS   3953..4267
FT   /product= "Omp A leader amino acid sequence"
FT   /tag= c
FT   CDS   4298..4363
FT   /product= "Kappa constant region amino acid sequence"
FT   /tag= d
FT   CDS   5179..6090
FT   /product= "pel B leader amino acid sequence"
FT   /tag= e
FT   CDS   /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT   sequence"

XX
PN   WO20024636-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46516.
XX
PR   08-DEC-2000; 2000US-254411P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;
XX
DR   WPI; 2002-537570/57.
DR   P-PSDB; AAO22535, AAO22536, AAO22537, AAO22538, AAO22539.
XX
PT   Novel plasmid useful in cloning and expression of foreign genetic
PT   information -
XX
PS   Claim 1; Fig 7A-E; 39pp; English.
XX
CC   The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC   6125, or 5683 base pairs fully defined in the specification. The
CC   invention more specifically relates to novel vectors capable of
CC   replication and expression of foreign genetic information in bacteria,
CC   such as, for example, cyanobacterium and E. coli. The new vectors have
CC   been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC   new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC   polynucleotide sequence represents the plasmid pRL5-CAT nucleic acid
CC   sequence of the invention.
XX
SQ   Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;

```
Query Match             100.0%;  Score 66;  DB 24;  Length 6122;
Best Local Similarity   100.0%;  Pred. No. 6.5e-14;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy         61 GCGGCC 66
              ||||||
Db       3110 GCGGCC 3115
```

RESULT 11
AAL41115
ID   AAL41115 standard; DNA; 6122 BP.
XX
AC   AAL41115;
XX
DT   16-OCT-2002 (first entry)
XX
DE   Plasmid pRL5 bsi-CAT nucleic acid sequence.
XX
KW   Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW   cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW   pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1013..1672
FT   /tag= a
FT   /product= "CAT amino acid sequence"
FT   CDS   3050..3115
FT   /tag= b
FT   CDS   3947..4267
FT   /product= "Omp A leader amino acid sequence"
FT   /tag= c
FT   CDS   4298..4363
FT   /product= "Kappa constant region amino acid sequence"

```
FT                /*tag= d
FT     CDS        /product= "pel B leader amino acid sequence"
FT                5179..6090
FT                /*tag= e
FT                /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT                sequence"
XX
PN     WO200246436-A2.
XX
PD     13-JUN-2002.
XX
PF     07-DEC-2001; 2001WO-US46516.
XX
PR     08-DEC-2000; 2000US-254411P.
XX
PA     (ALEX-) ALEXION PHARM INC.
XX
PI     Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;
XX
DR     WPI; 2002-537570/57.
DR     P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22541.
XX
PT     Novel plasmid useful in cloning and expression of foreign genetic
PT     information -
XX
PS     Disclosure; Fig 11A-E; 39pp; English.
XX
CC     The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC     6125, or 5683 base pairs fully defined in the specification. The
CC     invention more specifically relates to novel vectors capable of
CC     replication and expression of foreign genetic information in bacteria,
CC     such as, for example, cyanobacterium and E. coli. The new vectors have
CC     been designed to overcome certain drawbacks of the pComb3x plasmid. These
CC     new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC     polynucleotide sequence represents the plasmid pRL5 bsi-CAT nucleic acid
CC     sequence of the invention.
XX
SQ     Sequence 6122 BP; 1447 A; 1571 C; 1641 G; 1463 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 6122;
Best Local Similarity   100.0%;  Pred. No. 6.5e-14;
Matches    66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy     61 GCGGCC 66
          ||||||
Db   3110 GCGGCC 3115

RESULT 12
ABN84078
ID     ABN84078 standard; DNA; 6122 BP.
XX
AC     ABN84078;
XX
DT     23-SEP-2002  (first entry)
XX
DE     Plasmid pRL5-CAT.
XX
KW     Plasmid pTL5-CAT; vector; antibody; gene; ds.
XX
OS     Unidentified.
XX
FH     Key           Location/Qualifiers
FT     CDS           591..800
FT                   /*tag= a
FT                   /partial
FT                   /product= "Amp frag"
FT                   /note= "the CDS does not include a stop codon"
FT     CDS           1013..1672
```

```
FT
FT
FT     misc_feature  /*tag= b
FT                   /product= "Chloramphenicol transferase"
FT                   complement (2052..2651)
FT                   /*tag= c
FT                   /note= "ori"
FT     promoter      2831..3046
FT                   /*tag= d
FT                   /note= "lac promoter"
FT                   3009..3029
FT     misc_signal   /*tag= e
FT                   /note= "lac rep site"
FT     RBS           3036..3039
FT                   /*tag= f
FT     CDS           3050..3115
FT                   /*tag= g
FT                   /partial
FT                   /product= "OmpA leader"
FT                   /note= "the CDS does not include a stop codon"
FT     misc_feature  3121..3946
FT                   /*tag= h
FT                   /note= "light chain variable region stuffer"
FT     misc_feature  3951..4269
FT                   /*tag= i
FT                   /note= "Kappa Cns"
FT     CDS           4298..4363
FT                   /*tag= j
FT                   /partial
FT                   /product= "pelB leader"
FT                   /note= "the CDS does not include a stop codon"
FT     misc_feature  4365..5147
FT                   /*tag= k
FT                   /note= "heavy chain variable region stuffer"
FT     CDS           5485..5556
FT                   /*tag= l
FT                   /partial
FT                   /product= "Linker-His tag-HA tag"
FT                   /note= "the CDS does not include a start codon"
FT     CDS           5557..6090
FT                   /*tag= m
FT                   /partial
FT                   /gene= "Gene III"
FT                   /note= "the CDS does not include a start codon"
XX
PN     WO200246435-A2.
XX
PD     13-JUN-2002.
XX
PF     05-DEC-2001; 2001WO-US47452.
XX
PR     05-DEC-2000; 2000US-251440P.
XX
PA     (ALEX-) ALEXION PHARM INC.
XX
PI     Bowdish KS, Barbas-Frederickson S, Lin Y, Renshaw M, Wild M;
PI     McWhirter J;
XX
DR     WPI; 2002-537569/57.
XX
PT     Plasmid for in situ production of genes, comprises two template
PT     annealing sequences, such as downstream primer and upstream collar
PT     sequence and a restriction site located between the annealing sequences
PT     -
XX
PS     Disclosure; Fig 4A-T; 65pp; English.
XX
CC     The present sequence is that of plasmid pRL5-CAT, a derivative of
CC     pComb 3X which has been modified to contain chloramphenicol
CC     resistance. The invention relates to the engineering of plasmids,
CC     of which pRL-CAT is a preferred example, for in situ production of
CC     genes. It was found that nucleic acids encoding a polypeptide
CC     can be directly incorporated into a plasmid by DNA polymerisation
CC     or by reverse transcription of a nucleic acid template. Preferably,
CC     nucleic acids encoding at least a portion of an antibody can be
```

```
CC    directly incorporated into the plasmid by reverse transcription of
CC    mRNA. The plasmids are engineered to contain 2 template annealing
CC    sequences (see ABN84079-80 and ABN84082-83), i.e. a downstream
CC    primer that anneals to a first portion of a nucleic acid template,
CC    e.g. mRNA encoding at least a portion of an antibody, an upstream
CC    collar sequence that anneals to a second portion of the template,
CC    and at least 1 restriction site located between the 2 template
CC    annealing sequences. A single-stranded DNA plasmid vector is
CC    produced containing a nucleic acid encoding at least a portion of
CC    a polypeptide, e.g. a light chain and/or a heavy chain of an
CC    antibody. This vector can be transformed into a host cell and
CC    amplified.
XX
SQ    Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 6122;
Best Local Similarity   100.0%;  Pred. No. 6.5e-14;
Matches   66;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy       61 GCGGCC 66
            ||||||
Db     3110 GCGGCC 3115

RESULT 13
AAL41114
ID    AAL41114 standard; DNA; 6125 BP.
XX
AC    AAL41114;
XX
DT    16-OCT-2002 (first entry)
XX
DE    Plasmid pRL5 CAT-Asc nucleic acid sequence.
XX
KW    Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW    cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW    pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS    Unidentified.
XX
FH    Key             Location/Qualifiers
FT    CDS             1013..1672
FT                    /tag= a
FT                    /product= "CAT amino acid sequence"
FT    CDS             3050..3115
FT                    /tag= b
FT                    /product= "Omp A leader amino acid sequence"
FT    CDS             3953..4270
FT                    /tag= c
FT                    /product= "Kappa constant region amino acid sequence"
FT    CDS             4301..4366
FT                    /tag= d
FT                    /product= "pel B leader amino acid sequence"
FT    CDS             5182..6093
FT                    /tag= e
FT                    /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT                    sequence"
XX
PN    WO200246436-A2.
XX
PD    13-JUN-2002.
XX
PF    07-DEC-2001; 2001WO-US46516.
XX
PR    08-DEC-2000; 2000US-254411P.
XX
PA    (ALEX-) ALEXION PHARM INC.
XX
PI    Bowdish KS,  Barbas-frederickson S,  Wild M,  Mcwhirter J;
```

```
XX
DR    WPI; 2002-537570/57.
DR    P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22540.
XX
PT    Novel plasmid useful in cloning and expression of foreign genetic
PT    information -
XX
PS    Claim 1; Fig 9A-B; 39pp; English.
XX
CC    The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC    6125, or 5683 base pairs fully defined in the specification. The
CC    invention more specifically relates to novel vectors capable of
CC    replication and expression of foreign genetic information in bacteria,
CC    such as, for example, cyanobacterium and E. coli. The new vectors have
CC    been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC    new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC    polynucleotide sequence represents the plasmid pRL5 CAT-Asc nucleic acid
CC    sequence of the invention.
XX
SQ    Sequence 6125 BP; 1447 A; 1573 C; 1644 G; 1461 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 6125;
Best Local Similarity   100.0%;  Pred. No. 6.5e-14;
Matches   66;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy       61 GCGGCC 66
            ||||||
Db     3110 GCGGCC 3115

RESULT 14
AAQ92547
ID    AAQ92547 standard; DNA; 6166 BP.
XX
AC    AAQ92547;
XX
DT    11-MAR-1996 (first entry)
XX
DE    Expression vector, pPho-TT.
XX
KW    Human; Fab; variable chain; heavy; light; region; VH; VL; HIV; gp120;
KW    3b1; 3b3; 3b4; 3b9; MT4; humanised; monoclonal antibody; MAb;
KW    immunoreaction; neutralisation; passive immunotherapy; tetanus toxin;
KW    alkaline phosphatase; phoA; ss; cyclic.
XX
OS    Synthetic.
XX
PN    WO9511317-A1.
XX
PD    27-APR-1995.
XX
PF    19-OCT-1994;  94WO-US11907.
XX
PR    19-SEP-1994;  94US-0308841.
PR    19-OCT-1993;  93US-0139409.
PR    26-APR-1994;  94US-0233619.
XX
PA    (SCRI ) SCRIPPS RES INST.
XX
PI    Barbas CF,  Burton DR,  Lerner RA;
XX
DR    WPI; 1995-170235/22.
XX
PT    Synthetic human neutralising monoclonal antibodies to human
PT    immunodeficiency virus - used for diagnosis and immuno:therapy of
PT    HIV-induced disease
XX
PS    Example 2; Page 193-197; 249pp; English.
XX
```

CC   This sequence represents the expression vector, pPho-TT which is a
CC   modified version of the phagemid expression vector, pComb3 given in
CC   AAQ92546. pPho-TT provides for the expression of soluble Fabs which are
CC   secreted into the periplasmic space which is regulated from the alkaline
CC   phosphatase (phoA) promoter. This plasmid was used within the scope of
CC   the invention to express various mutagenised human Fab's which comprise
CC   heavy and light variable regions which bind to HIV gp120. pPho-TT
CC   consists of a DNA molecule having two cassettes to express two soluble
CC   proteins of a heavy chain and a light chain. The vector comprises,
CC   operatively linked 5' to 3', a first cassette consisting of the phoA
CC   promoter/operator sequences, an EcoRI restriction site, a ribosome
CC   binding site (RBS), an OmpA leader, a SfiI restriction site, a spacer
CC   region, a cloning region bordered by 5' sacI and 3' XbaI restriction
CC   sites, an NcoI restriction site between the two cassettes, and a second
CC   cassette consisting of an expression control RBS, a PelB leader, a human
CC   consensus amino terminus spacer region comprising the sequence EVQLLE,
CC   a cloning region bordered by 5' XhoI and 3' SpeI restriction sites
CC   followed by a SfiI site, expression control stop sequences and a NotI
CC   restriction site. The pPho-TT expression vector contains a light
CC   chain stuffer that is 1200 bp in length and a heavy chain stuffer that
CC   is 300 bp in length. The nucleotide sequences of the heavy and light
CC   chain stuffers encode the heavy and light chain variable domains of a
CC   tetanus toxin-specific Fab.
XX
SQ   Sequence 6166 BP; 1416 A; 1706 C; 1629 G; 1415 T; 0 other;

Query Match   100.0%;   Score 66; DB 16; Length 6166;
Best Local Similarity   100.0%;   Pred. No. 6.5e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4541 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 4600

Qy   61 GCGGCC 66
           ||||||
Db   4601 GCGGCC 4606

RESULT 15
AAQ10569
ID   AAQ10569 standard; DNA; 73 BP.
XX
AC   AAQ10569;
XX
DT   26-APR-1991 (first entry)
XX
DE   Omp A signal peptide-encoding sequence.
XX
KW   Mirabilis antiviral protein; MAP; Omp A signal sequence; ds.
XX
OS   Mirabilis jalapa.
XX
PN   EP414134-A.
XX
PD   27-FEB-1991.
XX
PF   16-AUG-1990;   90EP-0115718.
PR   17-AUG-1989;   89JP-0210767.
XX
PA   (NISB) JAPAN TOBACCO INC.
XX
PI   Habuka N, Akiyama K, Tsuge H, Matsumoto T, Noma M;
XX
DR   WPI; 1991-059386/09.
XX
PT   Gene encoding Mirabilis Antiviral Protien and OmpA signal peptide -
PT   for mass-prodn. of extracellular MAP.
XX
PS   Claim 1; page 8; 15pp; English.
XX
CC   This Omp A signal peptide-encoding sequence is contained in a

CC   recombinant plasmid, flanked on its 3' side by a Mirabilis
CC   antiviral protein (MAP) gene. The plasmid is used to transform
CC   E.coli host cells in the large-scale prodn. of MAP. The 5' end
CC   overhangs the 3' end of the complementary strand by TA and the 5'
CC   end of the complementary strand overhangs the 3' end of this sense
CC   strand by CTAG. See also AAQ10568.
XX
SQ   Sequence 73 BP; 16 A; 20 C; 21 G; 16 T; 0 other;

Query Match   91.2%;   Score 60.2; DB 12;   Length 73;
Best Local Similarity   95.4%;   Pred. No. 2.7e-12;
Matches   62; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db   2 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGCTACCGTAGCGCAG 61

Qy   61 GCGGCC 65
           || |||
Db   62 GCCGC 66

Search completed: August 18, 2003, 23:28:53
Job time : 24.2513 secs

THIS PAGE BLANK (USPTO)

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       August 18, 2003, 22:51:23 ; Search time 5.16534 Seconds
                                (without alignments)
                              5639.757 Million cell updates/sec

Title:        US-09-987-455-6
Perfect score: 66
Sequence:     1 atgaaaagacagctatcgc.........ctaccgtgcccaggcggcc 66

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     569978 seqs, 22069156b residues

Total number of hits satisfying chosen parameters:        1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents NA:*
              1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
              6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	66	100.0	3450	4	US-09-318-786-20	Sequence 20, Appl
2	66	100.0	4691	3	US-08-591-632-43/	Sequence 43, Appl
3	66	100.0	4691	4	US-09-611-451-43/	Sequence 43, Appl
4	66	100.0	6166	3	US-08-591-63z-51/	Sequence 51, Appl
5	66	100.0	6166	4	US-09-611-451-51/	Sequence 51, Appl
6	60.2	91.2	73	1	US-07-854-845B-2	Sequence 2, Appl
7	60.2	91.2	77	3	US-08-666-354A-10	Sequence 10, Appl
8	60.2	91.2	249	1	US-08-439-132-1	Sequence 1, Appl
9	60.2	91.2	249	6	5223407-1	Patent No. 5223407
10	60.2	91.2	548	3	US-08-828-741B-1	Sequence 1, Appli
11	60.2	91.2	548	4	US-09-160-567-1	Sequence 1, Appli
12	60.2	91.2	548	4	US-09-710-299-1	Sequence 1, Appli
13	59.6	90.3	458	3	US-08-880-829-21	Sequence 21, Appl
14	59.6	90.3	470	3	US-08-828-741B-10	Sequence 10, Appl
15	59.6	90.3	470	4	US-09-160-567-10	Sequence 10, Appl
16	59.6	90.3	470	4	US-09-710-299-10	Sequence 10, Appl
17	59.6	90.3	599	3	US-08-828-741B-7	Sequence 7, Appli
18	59.6	90.3	599	4	US-09-160-567-7	Sequence 7, Appli
19	59.6	90.3	599	4	US-09-710-299-7	Sequence 7, Appli
20	59.6	90.3	1031	3	US-08-828-741B-5	Sequence 5, Appli
21	59.6	90.3	1031	4	US-09-160-567-5	Sequence 5, Appli
22	59.6	90.3	1031	4	US-09-710-299-5	Sequence 5, Appli
23	59.6	90.3	1490	3	US-08-828-741B-3	Sequence 3, Appli
24	59.6	90.3	1490	4	US-09-160-567-3	Sequence 3, Appli
25	59.6	90.3	1490	4	US-09-710-299-3	Sequence 3, Appli
26	59.2	89.7	116	1	US-07-842-089E-23	Sequence 23, Appl
27	59.2	89.7	116	1	US-08-264-485-23	Sequence 23, Appl
28	59.2	89.7	245	6	5187153-24	Patent No. 5187153
29	59.2	89.7	245	4	5223482-26	Patent No. 5223482
30	59.2	89.7	1379	4	US-09-553-498-5	Sequence 5, Appli
31	59.2	89.7	1379	4	US-09-618-869-5	Sequence 5, Appli
32	59.2	89.7	1881	4	US-09-553-498-1	Sequence 1, Appli
33	59.2	89.7	1881	4	US-09-553-498-3	Sequence 3, Appli
34	59.2	89.7	1881	4	US-09-618-869-1	Sequence 1, Appli
35	59.2	89.7	1881	3	US-09-618-869-3	Sequence 3, Appli
36	59.2	89.7	4515	2	US-08-256-790-1	Sequence 1, Appli
37	58.8	89.1	63	1	US-07-842-089E-5	Sequence 5, Appli
38	58.8	89.1	63	2	US-08-264-485-5	Sequence 5, Appli
39	58.8	89.1	63	2	US-09-003-081-7	Sequence 7, Appli
40	58.8	89.1	63	3	US-08-648-262-7	Sequence 7, Appli
41	58.8	89.1	63	3	US-08-648-263-7	Sequence 7, Appli
42	58.8	89.1	63	4	US-09-904-196B-4	Sequence 4, Appli
43	58.8	89.1	63	4	US-09-230-233A-6	Sequence 6, Appli
44	58.8	89.1	103	1	US-08-215-138-12	Sequence 12, Appl
45	58.8	89.1	103	1	US-08-407-344-12	Sequence 12, Appl

```
                              ALIGNMENTS

RESULT 1
US-09-318-786-20
; Sequence 20, Application US/09318786
; Patent No. 6472147
; GENERAL INFORMATION:
;  APPLICANT: Janda, Kim D
;  APPLICANT: Wirsching, Peter
;  APPLICANT: Lerner, Richard A
;  APPLICANT: Gao, Changshou
;  TITLE OF INVENTION: METHODS FOR DISPLAY OF HETERODIMERIC PROTEINS ON
;  TITLE OF INVENTION: FILAMENTOUS PHAGE USING pVII AND pIX, COMPOSITIONS,
;  TITLE OF INVENTION: VECTORS AND COMBINATORIAL LIBRARIES
;  FILE REFERENCE: TSR0305S
;  CURRENT APPLICATION NUMBER: US/09/318,786
;  CURRENT FILING DATE: 1999-05-25
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;  LENGTH: 3450
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: Vector
US-09-318-786-20

  Query Match             100.0%;  Score 66;  DB 4;  Length 3450;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-13;
  Matches   66;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2667 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGCCCAG 2726

Qy         61 GCGGCC 66
              ||||||
Db       2727 GCGGCC 2732

RESULT 2
US-08-591-632-43
; Sequence 43, Application US/08591632
; Patent No. 6261558
; GENERAL INFORMATION:
;  APPLICANT: Barbas, Carlos F.
;  APPLICANT: Burton, Dennis R.
;  APPLICANT: Lerner, Richard A.
;  TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
;  TITLE OF INVENTION: ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
;  NUMBER OF SEQUENCES: 92
;  CORRESPONDENCE ADDRESS:
```

```
             ADDRESSEE:  The Scripps Research Institute, Office of
             ADDRESSEE:  Patent Counsel
             STREET:  10550 No. 6261558th Torrey Pines Road, TPC 8
             CITY:  La Jolla
             STATE:  CA
             COUNTRY:  USA
             ZIP:  92037
         COMPUTER READABLE FORM:
             MEDIUM TYPE:  Floppy disk
             COMPUTER:  IBM PC compatible
             OPERATING SYSTEM:  PC-DOS/MS-DOS
             SOFTWARE:  PatentIn Release #1.0
         CURRENT APPLICATION DATA:
             APPLICATION NUMBER:  US/08/591,632
             FILING DATE:  19-OCT-1994
         PRIOR APPLICATION DATA:
             APPLICATION NUMBER:  PCT/US94/11907
             FILING DATE:  19-OCT-1994
         PRIOR APPLICATION DATA:
             APPLICATION NUMBER:  US 08/308,841
             FILING DATE:  19-SEP-1994
         PRIOR APPLICATION DATA:
             APPLICATION NUMBER:  US 08/233,619
             FILING DATE:  26-APR-1994
         PRIOR APPLICATION DATA:
             APPLICATION NUMBER:  US 08/139,409
             FILING DATE:  19-OCT-1993
         ATTORNEY/AGENT INFORMATION:
             NAME:  Fitting, Thomas
             REGISTRATION NUMBER:  34,163
             REFERENCE/DOCKET NUMBER:  TSRI 332.3
         TELECOMMUNICATION INFORMATION:
             TELEPHONE:  (619) 784-2937
             TELEFAX:  (619) 784-9399
     INFORMATION FOR SEQ ID NO: 43:
         SEQUENCE CHARACTERISTICS:
             LENGTH:  4691 base pairs
             TYPE:  nucleic acid
             STRANDEDNESS:  single
             TOPOLOGY:  circular
         MOLECULE TYPE:  DNA (genomic)
US-08-591-632-43

Query Match          100.0%;  Score 66;  DB 3;  Length 4691;
Best Local Similarity 100.0%;  Pred. No. 1.1e-13;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676


RESULT 3
US-09-611-451-43
; Sequence 43, Application US/09611451
; Patent No. 6395275
; GENERAL INFORMATION:
         APPLICANT:  Barbas, Carlos F.
         APPLICANT:  Burton, Dennis R.
         APPLICANT:  Lerner, Richard A.
     TITLE OF INVENTION:  SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
                          ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
     NUMBER OF SEQUENCES:  92
     CORRESPONDENCE ADDRESS:
         ADDRESSEE:  The Scripps Research Institute, Office of
                     Patent Counsel
         STREET:  10550 No. 6395275th Torrey Pines Road, TPC 8
         CITY:  La Jolla
         STATE:  CA
```

```
             COUNTRY:  USA
             ZIP:  92037
         COMPUTER READABLE FORM:
             MEDIUM TYPE:  Floppy disk
             COMPUTER:  IBM PC compatible
             OPERATING SYSTEM:  PC-DOS/MS-DOS
             SOFTWARE:  PatentIn Release #1.0
         CURRENT APPLICATION DATA:
             APPLICATION NUMBER:  US/09/611,451
             FILING DATE:  06-Jul-2000
         PRIOR APPLICATION DATA:
             APPLICATION NUMBER:  08/591,632
             FILING DATE:  2001-10-29
             APPLICATION NUMBER:  US 08/308,841
             FILING DATE:  19-SEP-1994
             APPLICATION NUMBER:  US 08/233,619
             FILING DATE:  26-APR-1994
             APPLICATION NUMBER:  US 08/139,409
             FILING DATE:  19-OCT-1993
         ATTORNEY/AGENT INFORMATION:
             NAME:  Fitting, Thomas
             REGISTRATION NUMBER:  34,163
             REFERENCE/DOCKET NUMBER:  TSRI 332.3
         TELECOMMUNICATION INFORMATION:
             TELEPHONE:  (619) 784-2937
             TELEFAX:  (619) 784-9399
     INFORMATION FOR SEQ ID NO: 43:
         SEQUENCE CHARACTERISTICS:
             LENGTH:  4691 base pairs
             TYPE:  nucleic acid
             STRANDEDNESS:  single
             TOPOLOGY:  circular
         MOLECULE TYPE:  DNA (genomic)
         SEQUENCE DESCRIPTION:  SEQ ID NO: 43:
US-09-611-451-43

Query Match          100.0%;  Score 66;  DB 4;  Length 4691;
Best Local Similarity 100.0%;  Pred. No. 1.1e-13;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676


RESULT 4
US-08-591-632-51
; Sequence 51, Application US/0591632
; Patent No. 6261558
; GENERAL INFORMATION:
         APPLICANT:  Barbas, Carlos F.
         APPLICANT:  Burton, Dennis R.
         APPLICANT:  Lerner, Righard A.
     TITLE OF INVENTION:  SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
     TITLE OF INVENTION:  ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
     NUMBER OF SEQUENCES:  92
     CORRESPONDENCE ADDRESS:
         ADDRESSEE:  The Scripps Research Institute, Office of
         ADDRESSEE:  Patent Counsel
         STREET:  10550 No. 6261558th Torrey Pines Road, TPC 8
         CITY:  La Jolla
         STATE:  CA
         COUNTRY:  USA
         ZIP:  92037
     COMPUTER READABLE FORM:
         MEDIUM TYPE:  Floppy disk
         COMPUTER:  IBM PC compatible
         OPERATING SYSTEM:  PC-DOS/MS-DOS
         SOFTWARE:  PatentIn Release #1.0
```

```
;; CURRENT APPLICATION DATA:
;;    APPLICATION NUMBER: US/08/591,632
;;    FILING DATE: 19-OCT-1994
;; PRIOR APPLICATION DATA:
;;    APPLICATION NUMBER: PCT/US94/11907
;;    FILING DATE: 19-OCT-1994
;; PRIOR APPLICATION DATA:
;;    APPLICATION NUMBER: US 08/308,841
;;    FILING DATE: 19-SEP-1994
;; PRIOR APPLICATION DATA:
;;    APPLICATION NUMBER: US 08/233,619
;;    FILING DATE: 26-APR-1994
;; PRIOR APPLICATION DATA:
;;    APPLICATION NUMBER: US 08/139,409
;;    FILING DATE: 19-OCT-1993
;; ATTORNEY/AGENT INFORMATION:
;;    NAME: Fitting, Thomas
;;    REGISTRATION NUMBER: 34,163
;;    REFERENCE/DOCKET NUMBER: TSRI 332.3
;; TELECOMMUNICATION INFORMATION:
;;    TELEPHONE: (619) 784-2937
;;    TELEFAX: (619) 784-9399
;; INFORMATION FOR SEQ ID NO: 51:
;;    SEQUENCE CHARACTERISTICS:
;;       LENGTH: 6166 base pairs
;;       TYPE: nucleic acid
;;       STRANDEDNESS: single
;;       TOPOLOGY: circular
;;    MOLECULE TYPE: DNA (genomic)
US-08-591-632-51

Query Match            100.0%;  Score 66;  DB 3;  Length 6166;
Best Local Similarity  100.0%;  Pred. No. 1.2e-13;
Matches   66;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4541 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 4600

Qy    61 GCGGCC 66
          ||||||
Db  4601 GCGGCC 4606


RESULT 5
US-09-611-451-51
; Sequence 51, Application US/09611451
; Patent No. 6395275
; GENERAL INFORMATION:
;;    APPLICANT: Barbas, Carlos F.
;;               Burton, Dennis R.
;;               Lerner, Richard A.
;;    TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
;;                        ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
;;    NUMBER OF SEQUENCES: 92
;;    CORRESPONDENCE ADDRESS:
;;       ADDRESSEE: The Scripps Research Institute, Office of
;;                  Patent Counsel
;;       STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8
;;       CITY: La Jolla
;;       STATE: CA
;;       COUNTRY: USA
;;       ZIP: 92037
;;    COMPUTER READABLE FORM:
;;       MEDIUM TYPE: Floppy disk
;;       COMPUTER: IBM PC compatible
;;       OPERATING SYSTEM: PC-DOS/MS-DOS
;;       SOFTWARE: PatentIn Release #1.0
;;    CURRENT APPLICATION DATA:
;;       APPLICATION NUMBER: US/09/611,451
;;       FILING DATE: 06-Jul-2000
;;    PRIOR APPLICATION DATA:
;;       APPLICATION NUMBER: 08/591,632
```

```
;;       FILING DATE: 2001-10-29
;;       APPLICATION NUMBER: US 08/308,841
;;       FILING DATE: 19-SEP-1994
;;       APPLICATION NUMBER: US 08/233,619
;;       FILING DATE: 26-APR-1994
;;       APPLICATION NUMBER: US 08/139,409
;;       FILING DATE: 19-OCT-1993
;;    ATTORNEY/AGENT INFORMATION:
;;       NAME: Fitting, Thomas
;;       REGISTRATION NUMBER: 34,163
;;       REFERENCE/DOCKET NUMBER: TSRI 332.3
;;    TELECOMMUNICATION INFORMATION:
;;       TELEPHONE: (619) 784-2937
;;       TELEFAX: (619) 784-9399
;; INFORMATION FOR SEQ ID NO: 51:
;;    SEQUENCE CHARACTERISTICS:
;;       LENGTH: 6166 base pairs
;;       TYPE: nucleic acid
;;       STRANDEDNESS: single
;;       TOPOLOGY: circular
;;    MOLECULE TYPE: DNA (genomic)
;;    SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-611-451-51

Query Match            100.0%;  Score 66;  DB 4;  Length 6166;
Best Local Similarity  100.0%;  Pred. No. 1.2e-13;
Matches   66;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4541 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 4600

Qy    61 GCGGCC 66
          ||||||
Db  4601 GCGGCC 4606


RESULT 6
US-07-854-845B-2
; Sequence 2, Application US/07854845B
; Patent No. 5340732
; GENERAL INFORMATION:
;;    APPLICANT: HABUKA, No. 5340732iyuki
;;    APPLICANT: MIYANO, Masashi
;;    APPLICANT: MATSUMOTO, Takashi
;;    APPLICANT: NOMA, Masana
;;    TITLE OF INVENTION: ANTIVIRAL PROTEIN
;;    NUMBER OF SEQUENCES: 9
;;    CORRESPONDENCE ADDRESS:
;;       ADDRESSEE: BIRCH, STEWART, KOLASCH, & BIRCH
;;       STREET: 8110 Gatehouse Road, Suite 500 East
;;       CITY: Falls Church
;;       STATE: VA
;;       COUNTRY: USA
;;       ZIP: 22042
;;    COMPUTER READABLE FORM:
;;       MEDIUM TYPE: Floppy disk
;;       COMPUTER: IBM PC compatible
;;       OPERATING SYSTEM: PC-DOS/MS-DOS
;;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;;    CURRENT APPLICATION DATA:
;;       APPLICATION NUMBER: US/07/854,845B
;;       FILING DATE: 20-MAR-1992
;;       CLASSIFICATION: 530
;;    ATTORNEY/AGENT INFORMATION:
;;       NAME: Murphy Jr., Gerald M
;;       REGISTRATION NUMBER: 28,977
;;       REFERENCE/DOCKET NUMBER: 42-194P
;;    TELECOMMUNICATION INFORMATION:
;;       TELEPHONE: 703-205-8000
;;       TELEFAX: 703-205-8050
;;       TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 2:
```

```
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 73 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
US-07-854-845B-2

Query Match      91.2%;    Score 60.2;   DB 1;    Length 73;
Best Local Similarity   95.4%;   Pred. No. 4.9e-12;
Matches   62;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||| |||| ||| |||
Db    2 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 61

Qy   61 GCGGC 65
        || ||
Db   62 GCCGC 66

RESULT 7
US-08-666-354A-10
; Sequence 10, Application US/08666354A
; Patent No. 6040141
; GENERAL INFORMATION:
;   APPLICANT: KLAUSER, THOMAS
;   APPLICANT: KRAMER, JOACHIM
;   APPLICANT: MEYER, THOMAS F.
;   APPLICANT: POHLNER, JOHANNES
;   TITLE OF INVENTION: BACTERIA USED TO PRODUCE STABLE FUSION
;   TITLE OF INVENTION: PROTEINS AND METHOD FOR THEIR IDENTIFICATION
;   NUMBER OF SEQUENCES: 18
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP.
;     STREET: P.O. BOX 747
;     CITY: FALLS CHURCH
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22040-0747
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/666,354A
;     FILING DATE: 23-SEP-1996
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: SVENSSON, LEONARD R.
;     REGISTRATION NUMBER: 30,330
;     REFERENCE/DOCKET NUMBER: 147-157P
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-205-8000
;     TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 77 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: other nucleic acid
;   DESCRIPTION: /desc = "FRAGMENT OF PLASMID pJK165"
US-08-666-354A-10

Query Match      91.2%;    Score 60.2;   DB 3;    Length 77;
Best Local Similarity   95.4%;   Pred. No. 4.9e-12;
Matches   62;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60


Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||| |||| ||| |||
Qy   61 GCGGC 65
        || ||
Db   61 GCCGC 65

RESULT 8
US-08-439-132-1
; Sequence 1, Application US/08439132
; Patent No. 5646015
; GENERAL INFORMATION:
;   APPLICANT: Wong, W. K.R.
;   APPLICANT: Sutherland, Margaret L.
;   TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS FROM
;   TITLE OF INVENTION: E.COLI
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley & Lardner
;     STREET: Suite 500, 3000 K Street
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20007-5109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/439,132
;     FILING DATE: 11-MAY-1995
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Bent, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER: 16777/200/ALLE
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202) 672-5300
;     TELEFAX: (202) 672-5399
;     TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 249 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-439-132-1

Query Match      91.2%;    Score 60.2;   DB 1;    Length 249;
Best Local Similarity   95.4%;   Pred. No. 6.2e-12;
Matches   62;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||| |||| ||| |||
Db  102 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 161

Qy   61 GCGGC 65
        || ||
Db  162 GCCGC 166

RESULT 9
5223407-1
;Patent No. 5223407
; APPLICANT: WONG, RAYMOND W.K.;SUTHERLAND, MARGARET L.
; TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS
;FROM E.COLI
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/395,797
;   FILING DATE: 18-AUG-1989
```

```
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: 239,145
;        FILING DATE: 31-AUG-1988
;SEQ ID NO:1:
;     LENGTH: 249
5223407-1

Query Match             91.2%;    Score 60.2;  DB 6;  Length 249;
Best Local Similarity   95.4%;    Pred. No. 6.2e-12;
Matches    62;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db    102 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGGCAG 161

Qy     61 GCGGC 65
          || ||
Db    162 GCGGC 166

RESULT 10
US-08-828-741B-1
; Sequence 1, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
;     APPLICANT: Koentgen, Frank
;     APPLICANT: Suess, Gabriele M.
;     APPLICANT: Tarlinton, David M.
;     APPLICANT: Treutlein, Herbert R.
;     TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
;     TITLE OF INVENTION: PRODUCING SAME
;     NUMBER OF SEQUENCES: 14
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;        STREET: 400 Garden City Plaza
;        CITY: Garden City
;        STATE: New York
;        COUNTRY: United States of America
;        ZIP: 11530
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/828,741B
;        FILING DATE: 26-MAR-1997
;        CLASSIFICATION: 435
;     ATTORNEY/AGENT INFORMATION:
;        NAME: DiGiglio, Frank S.
;        REGISTRATION NUMBER: 31,346
;        REFERENCE/DOCKET NUMBER: 10591
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (516) 742-4343
;        TELEFAX: (516) 742-4366
;        TELEX: 230 901 SANS UR
;     INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 548 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     FEATURE:
;        NAME/KEY: CDS
;        LOCATION: 1..548
US-08-828-741B-1

Query Match             91.2%;    Score 60.2;  DB 3;  Length 548;
Best Local Similarity   95.4%;    Pred. No. 7.2e-12;
Matches    62;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGGCAG 60

Qy     61 GCGGC 65
          || ||
Db     61 GCGGC 65

RESULT 11
US-09-160-567-1
; Sequence 1, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
;     APPLICANT: Koentgen, Frank
;     APPLICANT: Suess, Gabriele M.
;     APPLICANT: Tarlinton, David M.
;     APPLICANT: Treutlein, Herbert R.
;     TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
;     TITLE OF INVENTION: PRODUCING SAME
;     NUMBER OF SEQUENCES: 14
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;        STREET: 400 Garden City Plaza
;        CITY: Garden City
;        STATE: New York
;        COUNTRY: United States of America
;        ZIP: 11530
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/09/160,567
;        FILING DATE:
;        CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: 08/828,741
;        FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;        NAME: DiGiglio, Frank S.
;        REGISTRATION NUMBER: 31,346
;        REFERENCE/DOCKET NUMBER: 10591
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (516) 742-4343
;        TELEFAX: (516) 742-4366
;        TELEX: 230 901 SANS UR
;     INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 548 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     FEATURE:
;        NAME/KEY: CDS
;        LOCATION: 1..548
US-09-160-567-1

Query Match             91.2%;    Score 60.2;  DB 4;  Length 548;
Best Local Similarity   95.4%;    Pred. No. 7.2e-12;
Matches    62;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGGCAG 60

Qy     61 GCGGC 65
          || ||
Db     61 GCGGC 65

RESULT 12
```

US-09-710-299-1
; Sequence 1, Application US/09710299
; Patent No. 6521741
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; Suess, Gabriele M.
; Tarlinton, David M.
; Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/710,299
; FILING DATE: 09-No. 6521741-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..548
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-710-299-1

Query Match   91.2%; Score 60.2; DB 4; Length 548;
Best Local Similarity   95.4%; Pred. No. 7.2e-12;
Matches   62; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTAGCGGCAG 60

Qy   61 GCGGC 65
          || ||
Db   61 GCCGC 65
```

RESULT 13
US-08-880-829-21
; Sequence 21, Application US/08880829
; Patent No. 5925559
; GENERAL INFORMATION:
; APPLICANT: Collins, John
; APPLICANT: Roettgen, Peter
; TITLE OF INVENTION: A Collection of Phagemids, A
; TITLE OF INVENTION: Collection of Escherichia Coli
; TITLE OF INVENTION: Cells Carrying The Phagemids, A
; TITLE OF INVENTION: Collection of Phagemid Particles
; TITLE OF INVENTION: Produced From Said Collection
; TITLE OF INVENTION: And Phagemid Particles
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Joseph T. Eisele
; ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,
; ADDRESSEE: Levy, Eisele and Richard
; STREET: 711 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017-4059
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/2" DISKETTE
; COMPUTER: IBM-XT COMPATIBLE
; OPERATING SYSTEM: DOS 3.3:
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,829
; FILING DATE: 23-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,668
; FILING DATE: 06/02/95
; APPLICATION NUMBER: German EP 94 108 689.4
; FILING DATE: 06/07/94
; ATTORNEY/AGENT INFORMATION:
; NAME: EISELE, JOSEPH T.
; REGISTRATION NUMBER: 25,331
; REFERENCE/DOCKET NUMBER: 2727-77
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 687-6000
; TELEFAX: (212) 682-3485
; TELEX: (212) 426767
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
US-08-880-829-21

Query Match   90.3%; Score 59.6; DB 2; Length 458;
Best Local Similarity   93.9%; Pred. No. 1.1e-11;
Matches   62; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  112 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTAGCGGCAG 171

Qy   61 GCGGCC 66
          || ||
Db  172 GCCGAC 177
```

RESULT 14

US-08-828-741B-10

; Sequence 10, Application US/08828741B

; Patent No. 6043069

; GENERAL INFORMATION:

; APPLICANT: Koentgen, Frank

; APPLICANT: Suess, Gabriele M.

; APPLICANT: Tarlinton, David M.

; APPLICANT: Treutlein, Herbert R.

; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF

; TITLE OF INVENTION: PRODUCING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/828,741B

; FILING DATE: 26-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: DiGiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 10591

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 470 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..470

US-08-828-741B-10

Query Match 90.3%; Score 59.6; DB 3; Length 470;
Best Local Similarity 93.9%; Pred. No. 1.1e-11;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||| ||||||||| || ||
Db 1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGCTGGCTACCGTAGCGGCAG 60

Qy 61 GCGGGCC 66
        |||| |—|
Db 61 GCCGAC 66

RESULT 15

US-09-160-567-10

; Sequence 10, Application US/09160567

; Patent No. 6326179

; GENERAL INFORMATION:

; APPLICANT: Koentgen, Frank

; APPLICANT: Suess, Gabriele M.

; APPLICANT: Tarlinton, David M.

; APPLICANT: Treutlein, Herbert R.

; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF

; TITLE OF INVENTION: PRODUCING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/160,567

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/828,741

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: DiGiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 10591

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 470 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..470

US-09-160-567-10

Query Match 90.3%; Score 59.6; DB 4; Length 470;
Best Local Similarity 93.9%; Pred. No. 1.1e-11;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||| ||||||||| || ||
Db 1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGCTGGCTACCGTAGCGGCAG 60

Qy 61 GCGGGCC 66
        |||| |—|
Db 61 GCCGAC 66

Search completed: August 19, 2003, 08:35:10
Job time : 5.16534 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 18.6243 Seconds
                                         (without alignments)
                                         7930.701 Million cell updates/sec

Title: US-09-987-455-6
Perfect score: 66
Sequence: 1 atgaaaaagacagctatcgc.........ctaccgtggccaggcggcc 66

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched: 150479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries

                SUMMARIES

Database :   Published Applications NA:*
  1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
  2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
  3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
  4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
  5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
  6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
  7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
  8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
  9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Match	Query Length	DB	ID	Description
1	66	100.0	66	11	US-09-987-457-1	Sequence 1, Appli
2	66	100.0	66	11	US-09-987-455-3	Sequence 3, Appli
3	66	100.0	66	11	US-09-987-455-6	Sequence 6, Appli
4	66	100.0	1128	11	US-09-987-455-2	Sequence 2, Appli
5	66	100.0	1128	11	US-09-987-455-5	Sequence 5, Appli
6	66	100.0	5149	14	US-10-006-593-60	Sequence 60, Appl
7	66	100.0	6122	14	US-10-006-591-1	Sequence 1, Appli
8	60.2	91.2	548	12	US-10-345-618-1	Sequence 1, Appli
9	59.6	90.3	470	12	US-10-345-618-10	Sequence 10, Appl
10	59.6	90.3	599	12	US-10-345-618-7	Sequence 7, Appli
11	59.6	90.3	932	9	US-09-809-517A-39	Sequence 39, Appl
12	59.6	90.3	1031	12	US-10-345-618-5	Sequence 5, Appli
13	59.6	90.3	1479	12	US-10-345-618-15	Sequence 15, Appl
14	59.6	90.3	1490	12	US-10-345-618-3	Sequence 3, Appli
15	59.6	90.3	1574	9	US-09-809-517A-38	Sequence 38, Appl
16	59.6	90.3	4425	9	US-09-809-517A-40	Sequence 40, Appl
17	59.2	89.7	4145	14	US-10-001-934-36	Sequence 36, Appl
18	59.2	89.7	5020	14	US-10-001-934-35	Sequence 35, Appl
19	59.2	89.7	5079	9	US-09-809-517A-41	Sequence 41, Appl
20	58.8	89.1	63	9	US-09-760-008A-4	Sequence 4, Appli
21	58.8	89.1	63	14	US-10-076-117-3	Sequence 3, Appli
22	58.8	89.1	63	14	US-10-192-294-4	Sequence 4, Appli
23	58.8	89.1	65	9	US-09-875-494-20	Sequence 20, Appl
24	58.8	89.1	108	11	US-09-848-616-8	Sequence 8, Appli
25	58.8	89.1	256	10	US-09-916-230-13	Sequence 13, Appli
26	58.8	89.1	256	11	US-09-848-616-18	Sequence 18, Appl
27	58.8	89.1	261	10	US-09-916-230-15	Sequence 15, Appl
28	58.8	89.1	261	11	US-09-848-616-20	Sequence 20, Appl
29	58.8	89.1	4614	9	US-09-912-165-17	Sequence 17, Appl
30	58.8	87.9	4657	9	US-09-912-165-18	Sequence 18, Appl
31	58	87.9	585	10	US-09-848-585-28	Sequence 28, Appl
32	58	87.9	585	10	US-09-848-585-30	Sequence 30, Appl
33	58	87.9	981	9	US-09-334-477-36	Sequence 36, Appl
34	58	87.9	990	9	US-09-334-477-38	Sequence 38, Appl
35	58	87.9	2321	10	US-09-995-396-2	Sequence 2, Appli
36	58	87.9	2337	10	US-09-995-396-3	Sequence 3, Appli
37	58	87.3	7083	10	US-09-995-396-1	Sequence 1, Appli
38	57.6	87.3	819	12	US-09-226-157-3	Sequence 3, Appli
39	57.2	86.7	903	14	US-10-033-399B-11	Sequence 11, Appl
40	55.4	83.9	93	13	US-10-004-832-3	Sequence 3, Appli
41	52.8	80.0	867	11	US-09-782-397-16	Sequence 16, Appl
c 42	52.8	80.0	867	11	US-09-782-397-18	Sequence 18, Appl
43	52.8	80.0	918	11	US-09-782-397-13	Sequence 13, Appl
c 44	52.8	80.0	918	11	US-09-782-397-15	Sequence 15, Appl
45	50	75.8	102	14	US-10-006-593-107	Sequence 107, App

                        ALIGNMENTS

RESULT 1
US-09-987-457-1
; Sequence 1, Application US/09987457
; Publication No. US2003001315OA1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-987-457-1

  Query Match   100.0%; Score 66; DB 11; Length 66;
  Best Local Similarity   100.0%; Pred. No. 4.6e-16;
  Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTCGCGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTCGCGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66

RESULT 2
US-09-987-455-3
; Sequence 3, Application US/09987455
; Publication No. US2003004972A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-987-455-3

Query Match   100.0%;   Score 66; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.6e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 3
US-09-987-455-6
; Sequence 6, Application US/09987455
; Publication No. US2003004972A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-987-455-6

Query Match   100.0%;   Score 66; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.6e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 4
US-09-987-455-2
; Sequence 2, Application US/09987455
; Publication No. US2003004972A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match   100.0%;   Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66

RESULT 5
US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US2003004972A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1128

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match   100.0%; Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches   66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
         ||||||
Db   61 GCGGCC 66

RESULT 6
US-10-006-593-60
; Sequence 60, Application US/10006593
; Publication No. US2003004973A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: vector
US-10-006-593-60

Query Match   100.0%; Score 66; DB 14; Length 5149;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches   66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2611 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
         ||||||
Db 2671 GCGGCC 2676

RESULT 7
US-10-006-591-1
; Sequence 1, Application US/10006591
; Publication No. US2003004973A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Renshaw, Mark
; APPLICANT: Wild, Martha
; APPLICANT: McWhirter, John
; TITLE OF INVENTION: ENGINEERED PLASMIDS AND THEIR USE FOR IN SITU PRODUCTION OF GENES

; FILE REFERENCE: 1087-3
; CURRENT APPLICATION NUMBER: US/10/006,591
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/251,440
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6122
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-10-006-591-1

Query Match   100.0%; Score 66; DB 14; Length 6122;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches   66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3050 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy   61 GCGGCC 66
         ||||||
Db 3110 GCGGCC 3115

RESULT 8
US-10-345-618-1
; Sequence 1, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(546)
; OTHER INFORMATION: Description of Artificial Sequence:LHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-1

Query Match   91.2%; Score 60.2; DB 12; Length 548;
Best Local Similarity 95.4%; Pred. No. 1.4e-13;
Matches   62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||| ||| |||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCCAG 60

Qy   61 GCGGC 65
         ||||||
Db   61 GCCGC 65

RESULT 9
US-10-345-618-10
; Sequence 10, Application US/1034518
; Publication No. US20030148484A1

```
; GENERAL INFORMATION:
;   APPLICANT: Koentgen, Frank
;   APPLICANT: Suess, Gabriele M.
;   APPLICANT: Tarlinton, David M.
;   APPLICANT: Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
;   FILE REFERENCE: 13474
;   CURRENT APPLICATION NUMBER: US/10/345,618
;   CURRENT FILING DATE: 2003-01-16
;   PRIOR APPLICATION NUMBER: US/09/509,031
;   PRIOR FILING DATE: 2000-06-09
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 470
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(468)
;   OTHER INFORMATION: Description of Artificial Sequence:Kappa
;   OTHER INFORMATION: nucleotide sequence
US-10-345-618-10

Query Match             90.3%;  Score 59.6;  DB 12;  Length 470;
Best Local Similarity   93.9%;  Pred. No. 2.3e-13;
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
             ||||||||||||||||||||||||||||| ||||||||||||||||| |||
Db         1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy        61 GCCGGCC 66
             ||||| |
Db        61 GCCGAC 66

RESULT 10
US-10-345-618-7
; Sequence 7, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
;   APPLICANT: Koentgen, Frank
;   APPLICANT: Suess, Gabriele M.
;   APPLICANT: Tarlinton, David M.
;   APPLICANT: Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
;   FILE REFERENCE: 13474
;   CURRENT APPLICATION NUMBER: US/10/345,618
;   CURRENT FILING DATE: 2003-01-16
;   PRIOR APPLICATION NUMBER: US/09/509,031
;   PRIOR FILING DATE: 2000-06-09
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(597)
;   OTHER INFORMATION: Description of Artificial Sequence:LHL.seq
;   OTHER INFORMATION: nucleotide sequence
US-10-345-618-7

Query Match             90.3%;  Score 59.6;  DB 12;  Length 599;
Best Local Similarity   93.9%;  Pred. No. 2.5e-13;
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
             ||||||||||||||||||||||||||||| ||||||||||||||||| |||
Db         1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60
```

```
Qy        61 GCCGGCC 66
             ||||| |
Db        61 GCCGAC 66

RESULT 11
US-09-809-517A-39
; Sequence 39, Application US/09809517A
; Patent No. US2002003473A1
; GENERAL INFORMATION:
;   APPLICANT: Lohning, Corinna
;   TITLE OF INVENTION: No. US20020034733A1el methods for displaying (poly)peptides/prote
;   TITLE OF INVENTION: particles via disulfide bonds
;   FILE REFERENCE: MORPHO/11
;   CURRENT APPLICATION NUMBER: US/09/809,517A
;   CURRENT FILING DATE: 2001-03-15
;   PRIOR APPLICATION NUMBER: EP 99114072.4
;   PRIOR FILING DATE: 1999-07-20
;   PRIOR APPLICATION NUMBER: EP 00103551.8
;   PRIOR FILING DATE: 2000-02-18
;   NUMBER OF SEQ ID NOS: 41
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
;   LENGTH: 932
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-39

Query Match             90.3%;  Score 59.6;  DB 9;  Length 932;
Best Local Similarity   93.9%;  Pred. No. 2.8e-13;
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
             ||||||||||||||||||||||||||||| ||||||||||||||||| |||
Db       266 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 325

Qy        61 GCCGGCC 66
             ||||| |
Db       326 GCCGAC 331

RESULT 12
US-10-345-618-5
; Sequence 5, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
;   APPLICANT: Koentgen, Frank
;   APPLICANT: Suess, Gabriele M.
;   APPLICANT: Tarlinton, David M.
;   APPLICANT: Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
;   FILE REFERENCE: 13474
;   CURRENT APPLICATION NUMBER: US/10/345,618
;   CURRENT FILING DATE: 2003-01-16
;   PRIOR APPLICATION NUMBER: US/09/509,031
;   PRIOR FILING DATE: 2000-06-09
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1031
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1029)
;   OTHER INFORMATION: Description of Artificial Sequence:TLHL nucleotide
;   OTHER INFORMATION: sequence
US-10-345-618-5

Query Match             90.3%;  Score 59.6;  DB 12;  Length 1031;
Best Local Similarity   93.9%;  Pred. No. 2.9e-13;
```

```
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||| ||||||||||||||||||||| || |||
Db      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy     61 GCGGCC 66
          |||  |
Db     61 GCCGAC 66


RESULT 13
US-10-345-618-15
; Sequence 15, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
;   APPLICANT:  Koentgen, Frank
;   APPLICANT:  Suess, Gabriele M.
;   APPLICANT:  Tarlinton, David M.
;   APPLICANT:  Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
;   FILE REFERENCE: 13474
;   CURRENT APPLICATION NUMBER: US/10/345,618
;   CURRENT FILING DATE:  2003-01-16
;   PRIOR APPLICATION NUMBER: US/09/509,031
;   PRIOR FILING DATE: 2000-06-09
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 1479
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:ccMTIgL
;   OTHER INFORMATION: nucleotide sequence
US-10-345-618-15

Query Match         90.3%;  Score 59.6;  DB 12;  Length 1479;
Best Local Similarity  93.9%;  Pred. No. 3.2e-13;
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||| ||||||||||||||||||||| || |||
Db     28 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 87

Qy     61 GCGGCC 66
          |||  |
Db     88 GCCGAC 93


RESULT 14
US-10-345-618-3
; Sequence 3. Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
;   APPLICANT:  Koentgen, Frank
;   APPLICANT:  Suess, Gabriele M.
;   APPLICANT:  Tarlinton, David M.
;   APPLICANT:  Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
;   FILE REFERENCE: 13474
;   CURRENT APPLICATION NUMBER: US/10/345,618
;   CURRENT FILING DATE:  2003-01-16
;   PRIOR APPLICATION NUMBER: US/09/509,031
;   PRIOR FILING DATE: 2000-06-09
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1490
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (1)..(1488)
;   OTHER INFORMATION: Description of Artificial Sequence:CATAB-TEV
;   OTHER INFORMATION: nucleotide sequence
US-10-345-618-3

Query Match         90.3%;  Score 59.6;  DB 12;  Length 1490;
Best Local Similarity  93.9%;  Pred. No. 3.2e-13;
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||| ||||||||||||||||||||| || |||
Db      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy     61 GCGGCC 66
          |||  |
Db     61 GCCGAC 66


RESULT 15
US-09-809-517A-38
; Sequence 38, Application US/09809517A
; Patent No. US20020034733A1
; GENERAL INFORMATION:
;   APPLICANT: Lohning, Corinna
;   TITLE OF INVENTION: No. US20020034733A1el methods for displaying (poly)peptides/prote
;   TITLE OF INVENTION: particles via disulfide bonds
;   FILE REFERENCE: MORPHO/11
;   CURRENT APPLICATION NUMBER: US/09/809,517A
;   CURRENT FILING DATE:  2001-03-15
;   PRIOR APPLICATION NUMBER: EP 99114072.4
;   PRIOR FILING DATE: 1999-07-20
;   PRIOR APPLICATION NUMBER: EP 00103551.8
;   PRIOR FILING DATE: 2000-02-18
;   NUMBER OF SEQ ID NOS: 41
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
;   LENGTH: 1574
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-38

Query Match         90.3%;  Score 59.6;  DB 9;  Length 1574;
Best Local Similarity  93.9%;  Pred. No. 3.2e-13;
Matches    62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||| ||||||||||||||||||||| || |||
Db    266 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 325

Qy     61 GCGGCC 66
          |||  |
Db    326 GCCGAC 331


Search completed: August 19, 2003, 14:22:28
Job time : 19.6243 secs
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
(c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 22:44:33 ; Search time 193.082 Seconds
                                            (without alignments)
                                8307.845 Million cell updates/sec

Title: US-09-987-455-6
Perfect score: 66
Sequence: 1 atgaaaagacagctatcgc.........ctaccgtggcccaggcggcc 66

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215223B056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
               1: em_estba:*
               2: em_esthum:*
               3: em_estin:*
               4: em_estmu:*
               5: em_estcov:*
               6: em_estpl:*
               7: em_estro:*
               8: em_htc:*
               9: gb_est1:*
              10: gb_est2:*
              11: gb_htc:*
              12: gb_est3:*
              13: gb_est4:*
              14: gb_est5:*
              15: em_estfun:*
              16: em_estom:*
              17: em_gss_hum:*
              18: em_gss_inv:*
              19: em_gss_pln:*
              20: em_gss_vrt:*
              21: em_gss_fun:*
              22: em_gss_mam:*
              23: em_gss_mus:*
              24: em_gss_pro:*
              25: em_gss_rod:*
              26: em_gss_phg:*
              27: em_gss_vrl:*
              28: gb_gss1:*
              29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	48	72.7	655	28	AF075794	AF075794 AF075794
c 2	27.8	42.1	386	14	CB806580	CB806580 AMGNNUC:S
c 3	26.8	40.6	310	9	AI635506	AI635506 ts95a03.x
c 4	26.8	40.6	312	9	AW004057	AW004057 wq85e09.x

c 5	26.8	40.6	321	10	BB503913	BB503913 h35e09.x
c 6	26.8	40.6	325	9	AA398159	AA398159 zt60e09.s
7	26.8	40.6	340	9	AA401699	AA401699 zt60e09.r
c 8	26.8	40.6	409	13	BY641012	BY641012 BY641012
9	26.8	40.6	569	13	BU783334	BU783334 in02g05.y
10	26.8	40.6	667	13	BX099539	BX099539 BX099539
c 11	26.8	40.6	694	29	AL301311	AL301311 Tetraodon
12	26.8	40.6	786	12	BI560274	BI560274 60325351l
13	26.8	40.6	871	13	BU853590	BU853590 AGENCOURT
c 14	26.6	40.3	268	28	AQ905715	AQ905715 GSSTc0708
15	26.6	40.3	359	9	AU244728	AU244728 AU244728
16	26.6	40.3	377	9	AU244560	AU244560 AU244560
17	26.6	40.3	450	10	BG234261	BG234261 da39a07.
c 18	26.6	40.3	472	28	AQ904535	AQ904535 GSSTc0415
19	26.6	40.3	612	12	BJ031326	BJ031326 BJ031326
20	26.6	40.3	645	13	BQ400448	BQ400448 NISC_mp09
21	26.6	40.3	753	14	CA808042	CA808042 CA12Li0lI
22	26.6	40.3	833	14	CB349272	CB349272 CAB2SG000
23	26.6	40.3	885	13	BQ737041	BQ737041 AGENCOURT
24	26.6	40.3	888	13	BQ737333	BQ737333 AGENCOURT
25	26.6	40.3	889	14	CD101043	CD101043 AGENCOURT
26	26.6	40.3	897	13	BQ733411	BQ733411 AGENCOURT
27	26.6	40.3	980	14	BQ206009	BQ206009 AGENCOURT
28	26.6	40.3	1124	14	CB206058	CB206058 AGENCOURT
29	26.2	39.7	271	9	AU184837	AU184837 AU184837
c 30	26.2	39.7	374	13	BQ094120	BQ094120 san42d04.
c 31	26.2	39.7	695	14	CD355744	CD355744 UI-M-FY0-
c 32	26.2	39.7	973	13	BQ938965	BQ938965 AGENCOURT
33	26	39.4	495	28	AQ833969	AQ833969 HS_5261_A
34	26	39.4	853	29	AL260159	AL260159 Tetraodon
c 35	26	39.4	950	29	AL292142	AL292142 Tetraodon
c 36	26	39.4	1016	29	AL215785	AL215785 Tetraodon
37	26	39.4	1387	29	B2557231	B2557231 pacsl-60
38	25.6	38.8	291	14	H33495	H33495 EST109553 R
39	25.6	38.8	331	14	CA377775	CA377775 656425 NC
40	25.6	38.8	438	14	CA373964	CA373964 648233 NC
41	25.6	38.8	621	9	AL966295	AL966295 AL966295
42	25.6	38.8	652	9	AL644617	AL644617 AL644617
43	25.6	38.8	673	9	AL655729	AL655729 AL655729
c 44	25.6	38.8	687	14	CA369617	CA369617 646096 NC
45	25.6	38.8	691	9	AL630991	AL630991 AL630991

ALIGNMENTS

RESULT 1
AF075794/c
LOCUS   AF075794   655 bp   DNA   linear   GSS 29-AUG-2000
DEFINITION   AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella
             typhimurium genomic clone 107-T3, genomic survey sequence.
ACCESSION   AF075794
VERSION   AF075794.1 GI:3320664
KEYWORDS   GSS.
SOURCE   Salmonella typhimurium
  ORGANISM   Salmonella typhimurium
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Salmonella.
REFERENCE   1 (bases 1 to 655)
  AUTHORS   Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.
  TITLE   Sample sequencing of a Salmonella typhimurium LT2 lambda library:
             comparison to the Escherichia coli K12 genome
  JOURNAL   FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
  MEDLINE   99243757
   PUBMED   10227170
COMMENT   Contact: McClelland M
             Molecular Biology
             Sidney Kimmel Cancer Center
             3099 Science Park Road, San Diego, CA 92121, USA
             Email: mcclelland@lifsci.sdsu.edu
             Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..655

Right column

```
AI635506/c
LOCUS       AI635506                310 bp    mRNA    linear   EST 14-DEC-1999
DEFINITION  te95a03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2238988 3',
            mRNA sequence.
ACCESSION   AI635506
VERSION     AI635506.1  GI:4688836
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 310)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 383   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES    Location/Qualifiers
     source 1..310
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2238988"
            /tissue_type="pooled germ cell tumors"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_GC6"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
            polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
            from the normalized library NCI_CGAP_GC4 was prepared, and
            ss circles were made in vitro. Following HAP purification,
            this DNA was used as tracer in a subtractive hybridization
            reaction. The driver was PCR-amplified cDNAs from a pool
            of 5,000 clones made from the same library (cloneIDs
            1257096-1258631, 1469064-1470983, and 1475592-1476743).
            Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT   71 a    72 c    64 g   103 t
ORIGIN

Query Match            40.6%;   Score 26.8; DB 9; Length 310;
Best Local Similarity  64.5%;   Pred. No. 97;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy    5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCCG 64
          |||||||||| |||| |||||| | ||||| || |||||||||| |||||
Db  268 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTGTTAGCAGAGCCCAGGCAG 209

Qy   65 CC 66
          ||
Db  208 CC 207

RESULT 4
AW004057/c
LOCUS       AW004057                312 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  wq85e09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2478856 3',
            mRNA sequence.
ACCESSION   AW004057
VERSION     AW004057.1  GI:5850973
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
```

Left column

```
            /organism="Salmonella typhimurium"
            /mol_type="genomic DNA"
            /strain="LT2"
            /db_xref="taxon:602"
            /clone="107-T3"
            /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
            /note="Vector: Lambda DASH II; sequenced using Li-Cor
            sequencer"
BASE COUNT  195 a   158 c   139 g   161 t     2 others
ORIGIN

Query Match            72.7%;   Score 48; DB 28; Length 655;
Best Local Similarity  85.2%;   Pred. No. 8.9e-06;
Matches 52; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy    5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCCG 64
          : |||||||||||||||||||||||||||||||||||||| || |||||
Db  298 AAAAAGACNBTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCG 239

Qy   65 C 65
          |
Db  238 C 238

RESULT 2
CB806580/c
LOCUS       CB806580                386 bp    mRNA    linear   EST 16-MAY-2003
            AMGNNUC:SRPG2-00054-G9-A srpg2 (10238) Rattus norvegicus cDNA clone
DEFINITION  srpg2-00054-g9 5', mRNA sequence.
ACCESSION   CB806580
VERSION     CB806580.1  GI:29924779
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 386)
AUTHORS     Amgen EST Program.
TITLE       Amgen Rat EST Program.
JOURNAL     Unpublished
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00054 row: g column: 9.
FEATURES    Location/Qualifiers
     source 1..386
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /clone="srpg2-00054-g9"
            /tissue_type="peneal gland brain"
            /clone_lib="srpg2 (10238)"
            /note="Vector: pSPORT1; Site_1: Sall; Site_2: NotI; peneal
            gland brain region"
BASE COUNT   57 a   149 c   119 g    60 t     1 others
ORIGIN

Query Match            42.1%;   Score 27.8; DB 14; Length 386;
Best Local Similarity  64.1%;   Pred. No. 48;
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy    2 TGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGG 61
          ||||| |||| |||| ||| |  ||||| || |||||||||  |||||
Db   85 TGAAGACGGCAGCGCGCGGCGGCGCGCCAGCCGGGCCGCGGCGCTAGGNCAGCGACGGCGGCCCCGG 26

Qy   62 CGGC 65
          ||||
Db   25 CGGC 22

RESULT 3
```

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

R. Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 313.

FEATURES
  source

Location/Qualifiers
1..321
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:3209992"
  /tissue_type="pooled germ cell tumors"
  /lab_host="DH10B"
  /clone_lib="NCI_CGAP_GC6"
  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT
ORIGIN

75 a   77 c   65 g   104 t

Query Match   40.6%;   Score 26.8;   DB 10;   Length 321;
Best Local Similarity 64.5%;   Pred. No. 98;
Matches 40;   Conservative   0;   Mismatches   22;   Indels   0;   Gaps   0;

Qy   5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
       ||||| |||||| |||| ||| ||| |||||||||||||| ||| ||||||||||||||
Db   269 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 210

Qy   65 CC 66
       ||
Db   209 CC 208

RESULT 6
AA398159/c
LOCUS   AA398159   325 bp   mRNA   linear   EST 12-AUG-1997
DEFINITION zt60e09.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726760
  3', mRNA sequence.
ACCESSION   AA398159
VERSION   AA398159.1 GI:2051405
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 325)
AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
  Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
  , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
  ,T., Waterston,R. and Wilson,R.
TITLE   WashU-Merck EST Project 1997
JOURNAL   Unpublished
COMMENT   Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 392   Std Error: 0.00
  Seq primer: -41m13 fwd. ET from Amersham.

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 312)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES
  source

Location/Qualifiers
1..312
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:2478856"
  /tissue_type="pooled germ cell tumors"
  /lab_host="DH10B"
  /clone_lib="NCI_CGAP_GC6"
  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT
ORIGIN

72 a   72 c   64 g   104 t

Query Match   40.6%;   Score 26.8;   DB 9;   Length 312;
Best Local Similarity 64.5%;   Pred. No. 97;
Matches 40;   Conservative   0;   Mismatches   22;   Indels   0;   Gaps   0;

Qy   5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
       ||||| |||||| |||| ||| ||| |||||||||||||| ||| ||||||||||||||
Db   269 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 210

Qy   65 CC 66
       ||
Db   209 CC 208

RESULT 5
BE503913/c
LOCUS   BE503913   321 bp   mRNA   linear   EST 04-AUG-2000
  hz35e09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3209992 3',
  mRNA sequence.
ACCESSION   BE503913
VERSION   BE503913.1 GI:9706321
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 321)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

```
FEATURES             Location/Qualifiers
     source          1..325
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:5923670"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:726760"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares testis NHT"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc. and primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT           72 c    64 g   117 t
ORIGIN

Query Match             40.6%;  Score 26.8;  DB 9;  Length 325;
Best Local Similarity   64.5%;  Pred. No. 98;
Matches   40;  Conservative   0;  Mismatches   22;  Indels   0;  Gaps   0;

Qy           5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
               |||| ||||| |||| |  ||| ||  |||| ||  |  |||| ||    ||||||||
Db         282 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 223

Qy          65 CC 66
               ||
Db         222 CC 221

RESULT 7
AA401699
LOCUS       AA401699               340 bp    mRNA    linear   EST 12-AUG-1997
DEFINITION  zt60e09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726760
            5', mRNA sequence.
ACCESSION   AA401699
VERSION     AA401699.1  GI:2057290
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 340)
AUTHORS     Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
TITLE       WashU-Merck EST Project 1997
JOURNAL     Unpublished
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 392   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..340
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:5923670"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:726760"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares testis NHT"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc. and primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT          106 a    75 c    82 g    77 t
ORIGIN

Query Match             40.6%;  Score 26.8;  DB 9;  Length 340;
Best Local Similarity   64.5%;  Pred. No. 1e+02;
Matches   40;  Conservative   0;  Mismatches   22;  Indels   0;  Gaps   0;

Qy           5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
               |||| ||||| |||| |  ||| ||  |||| ||  |  |||| ||    ||||||||
Db          76 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 135

Qy          65 CC 66
               ||
Db         136 CC 137

RESULT 8
BY641012/c
LOCUS       BY641012               409 bp    mRNA    linear   EST 15-DEC-2002
DEFINITION  BY641012 RIKEN full-length enriched, visual cortex Mus musculus
            cDNA clone K430341N20 3', mRNA sequence.
ACCESSION   BY641012
VERSION     BY641012.1  GI:26976194
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 409)
AUTHORS     Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.
            ,Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.
            ,Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,
            Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
            Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani
            ,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
            ,A., Frazer,K.S., Gaasterland,T., Gariboldi,C., Gissi,C., Godzik,A.
            ,Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
            Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
            King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
            ,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
            ,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
            Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
            Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
            ,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
            ,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
            ,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,
            Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa
            ,M., Yang,Y., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
            Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
            ,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.
            ,Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
            ,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata
            ,Y., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
            ,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE       Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL     Nature 420, 563-573 (2002)
MEDLINE     22354683
PUBMED      12466851
COMMENT     Contact: Yoshihide Hayashizaki
```

REFERENCE   1 (bases 1 to 569)
AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
             Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
             Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
             Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
             ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
             , Jackson,Y. and Bowers,Y.
TITLE   Endocrine Pancreas Consortium
JOURNAL   Unpublished
COMMENT   Other ESTs: in02g05.x1
             Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
             Endocrine Pancreas Consortium
             Harvard University, Howard Hughes Medical Institute
             Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
             MA 02138
             Tel: 617-495-1812
             Fax: 617-495-8557
             Email: dmelton@biohp.harvard.edu
             Library was constructed by Dr. J. Ferrer In vivo mass-excised to
             pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
             University Genome Sequencing Center For information on obtaining a
             clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
             Seq primer: -40RP from Gibco
             High quality sequence stop: 429.
FEATURES   Location/Qualifiers
     source   1..569
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:6123441"
                 /tissue_type="insulinoma"
                 /lab_host="DH10B (phage-resistant)"
                 /clone_lib="Human insulinoma"
                 /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                 XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                 pBluescript SK- by Dr. H. Inoue following the Washington
                 University protocol
                 (http://genome.wustl.edu/est/lambda_protocol.shtml).
                 please contact Hiroshi Inoue, MD/PhD for further
                 information on this library (Metabolism Division, Permutt
                 Laboratory, Washington University School of Medicine, Box
                 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                 is a Washington University Pancreas EST project library. "
BASE COUNT   118 a   160 c   144 g   147 t
ORIGIN

Query Match   40.6%; Score 26.8; DB 13; Length 569;
Best Local Similarity   64.5%; Pred. No. 1.2e+02;
Matches   40; Conservative   0; Mismatches   22; Indels   0; Gaps   0;

Qy   5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCCGCTACCGTGGCCCAGGCGG 64
              ||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   253 AAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTGTTAGCAGACCCCAGGCAG 312

Qy   65 CC 66
              ||
Db   313 CC 314

RESULT 10
BX099539
LOCUS   BX099539
DEFINITION BX099539 Soares_testis_NHT Homo sapiens cDNA clone IMAGp998G231928
            ; IMAGE:782638, mRNA sequence.
ACCESSION   BX099539
VERSION   BX099539.1 GI:27844114
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 667)

Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane
,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A.
, Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K.
, Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami
,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michela Fagiolini and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa,Wako-shi,Saitama 351-0198 Japan) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES   Location/Qualifiers
     source   1..409
                 /organism="Mus musculus"
                 /mol_type="mRNA"
                 /strain="C57BL/6J"
                 /db_xref="taxon:10090"
                 /clone="K43034N20"
                 /tissue_type="visual cortex"
                 /clone_lib="RIKEN full-length enriched, visual cortex"
BASE COUNT   89 a   96 c   87 g   135 t   2 others
ORIGIN

Query Match   40.6%; Score 26.8; DB 13; Length 409;
Best Local Similarity   64.5%; Pred. No. 1.1e+02;
Matches   40; Conservative   0; Mismatches   22; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              | |||||||||||||||||||||||||||||||||||| ||||| |||||||||||||||
Db   102 AGGAAAAAGAAGGCAAGCAGGGTTAGAGAGGTGGTGGCTGGGTTCTCAAGCTTTCCACAG 43

Qy   61 GC 62
              ||
Db   42 GC 41

RESULT 9
BU783334
LOCUS   BU783334   569 bp   mRNA   linear   EST 11-OCT-2002
DEFINITION in02g05.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123441
            5', mRNA sequence.
ACCESSION   BU783334
VERSION   BU783334.1 GI:23827413
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

            BX099539   667 bp   mRNA   linear   EST 06-FEB-2003

```
AUTHORS     Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
TITLE       Human UnigeneSet - RZPD3
JOURNAL     Unpublished
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998G231928.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/cloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from R2PD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES    Location/Qualifiers
     source  1..667
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGp998G231928 ; IMAGE:782638"
            /sex="male"
            /lab_host="DH10B"
            /clone_lib="Soares testis NHT"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
            polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from Clontech Laboratories
            , Inc., and primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCGGCCGCCCAATTTTTTTTTTTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT7T3 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT  195 a    142 c    167 g    162 t    1 others
ORIGIN

Query Match         40.6%;    Score 26.8;  DB 13;   Length 667;
Best Local Similarity  64.5%;  Pred. No. 1.2e+02;
Matches         40; Conservative   0; Mismatches   22;  Indels    0;  Gaps    0;

Qy     5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
           ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    43 AAAAAACAGCCATCACCAGGCCATTCAGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 102

Qy    65 CC 66
           ||
-Db   103 CC 104

RESULT 11
CNS04PHI/c
LOCUS       CNS04PHI          694 bp    DNA      linear    GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            127D11 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL301311
VERSION     AL301311.1  GI:8178354
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
```

```
TITLE       Saurin,W. and Weissenbach,J.
            Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
JOURNAL     Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE     20296633
PUBMED      10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Bilault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
  PUBMED    10899143
REFERENCE   3   (bases 1 to 694)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, Please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES    Location/Qualifiers
     source  1..694
            /organism="Tetraodon nigroviridis"
            /mol_type="genomic DNA"
            /db_xref="taxon:99883"
            /clone="127D11"
            /clone_lib="G"
            /note="Genoscope sequence ID : COBG127CB06SP1-end :
            PUC-Ori"
BASE COUNT  189 a    122 c    186 g    175 t    22 others
ORIGIN

Query Match         40.6%;    Score 26.8;  DB 29;   Length 694;
Best Local Similarity  60.6%;  Pred. No. 1.3e+02;
Matches         40; Conservative   2; Mismatches   24;  Indels    0;  Gaps    0;

Qy     1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
           | |||||| ||  ||| || ||||| | |||||||||||  :|  | ||||||||||||
Db   278 ACGTKAATGGCGGCTTTTGCGAARGCTCTKACAATACAKGAAGACGGTATCGTGGCCCAG 219

Qy    61 GCGGCC 66
           ||| ||
Db   218 GGCTCC 213

RESULT 12
BI560274
LOCUS       BI560274              786 bp    mRNA     linear    EST 05-SEP-2001
DEFINITION  603253511F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295746 5',
            mRNA sequence.
ACCESSION   BI560274
VERSION     BI560274.1  GI:15447588
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11748 row: f column: 03
High quality sequence stop: 772.
```
FEATURES        Location/Qualifiers
source          1..786
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5295746"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_97"
                /note="Organ: testis; Vector: pBluescriptR (modified
                pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTTVN-3',
                size-selected for average insert size 2.2 kb and
                normalized to ROT 5. This is a primary library enriched
                for full-length clones and constructed using the
                Cap-trapper method (Carninci, in preparation). Library
                constructed by M. Brownstein (NIMH/NHGRI, National
                Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      238 a    168 c    187 g    193 t
ORIGIN
```
Query Match   40.6%;   Score 26.8; DB 12; Length 786;
Best Local Similarity   64.5%;   Pred. No. 1.3e+02;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy   5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
        ||| |||| || || |  | ||| || | |||| | ||| | |||||||||
Db  56 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 115

Qy  65 CC 66
        ||
Db 116 CC 117
```

RESULT 13
BU853590
```
LOCUS       BU853590                871 bp    mRNA    linear    EST 16-OCT-2002
DEFINITION  AGENCOURT_10418173 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6620518
            5', mRNA sequence.
ACCESSION   BU853590
VERSION     BU853590.1  GI:24038556
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 871)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2870 row: 1 column: 22
            High quality sequence stop: 454.
FEATURES        Location/Qualifiers
source          1..871
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6620518"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_lib="NIH_MGC_82"
```

                /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                3' adaptors were used in cloning as follows: 5' adaptor
                sequence: 5'-CACGGCCATTATGGCC-3', and 3' adaptor sequence:
                5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                C, or G and N = A, C, G, or T). Average insert size
                1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
                inserts by PCR. This library was enriched for full-length
                clones and was constructed by Clontech Laboratories (Palo
                Alto, CA)."
```
BASE COUNT      240 a    226 c    203 g    202 t
ORIGIN
```
Query Match   40.6%;   Score 26.8; DB 13; Length 871;
Best Local Similarity   64.5%;   Pred. No. 1.4e+02;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy   5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
        ||| |||| || || |  | ||| || | |||| | ||| | |||||||||
Db  38 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 97

Qy  65 CC 66
        ||
Db  98 CC 99
```

RESULT 14
AQ905715/c
```
LOCUS       AQ905715              268 bp    DNA    linear    GSS 09-JAN-2001
DEFINITION  GSSTc07088 Trypanosoma cruzi random genomic library Trypanosoma
            cruzi genomic clone G55P8, genomic survey sequence.
ACCESSION   AQ905715
VERSION     AQ905715.3  GI:10139418
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi
ORGANISM    Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1 (bases 1 to 268)
AUTHORS     Aguero,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE       A random sequencing approach for the analysis of the trypanosoma
            cruzi genome: general structure, large gene and repetitive DNA
            families, and gene discovery
JOURNAL     Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE     20568489
PUBMED      11116094
COMMENT     On Sep 14, 2000 this sequence version replaced gi:9373324.
            Contact: Sanchez D.O.
            Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
            San Martin)
            Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos
            Aires, Argentina
            Tel: (54-11) 4580/7255/7
            Fax: (54-11) 4752-9639
            Email: dsanchez@iib.unsam.edu.ar
            Sequences were basecalled with phred and vector was masked with
            crossmatch (see http://genome.washington.edu). Sequences were then
            trimmed from both ends to remove low quality bases and masked
            vector.
            Seq primer: T7
            Class: shotgun.
FEATURES        Location/Qualifiers
source          1..268
                /organism="Trypanosoma cruzi"
                /mol_type="genomic DNA"
                /strain="CL-Brener"
                /db_xref="taxon:5693"
                /clone="G55P8"
                /cell_type="epimastigote"
                /clone_lib="Trypanosoma cruzi random genomic library"
                /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
                randomly sheared using a nebulizer and the 1 to 2 Kb range
                was gel purified and cloned into the dephosphoryated
```

```
BASE COUNT        60 a        63 c        84 g        61 t
ORIGIN                       HincII site of the vector"

Query Match         40.3%;   Score 26.6;  DB 28;   Length 268;
Best Local Similarity  71.4%;  Pred. No. 1.1e+02;
Matches    35;  Conservative   0;  Mismatches  14;  Indels   0;  Gaps   0;

Qy     16  ATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG  64
           |||||||| |||| ||||||  |||||| |||| |||| |||| |||||
Db    189  ATCGTGCTGGCTGTGGGCGCGGGTGCGCTTTGGCTCCCGTCGCCCTGGCGG  141


RESULT 15
AU244728
LOCUS       AU244728                 359 bp     mRNA        linear   EST  21-FEB-2002
DEFINITION  AU244728 Shibata Xenopus AEM lambda-ZAP II cDNA library Xenopus
            laevis cDNA clone p5el2 5', mRNA sequence.
ACCESSION   AU244728
VERSION     AU244728.1  GI:18850654
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 359)
  AUTHORS   Shibata,M., Itoh,M., Ohmori,S., Shinga,J. and Taira,M.
  TITLE     Systematic screening and expression analysis of the head organizer
            genes in Xenopus embryos
  JOURNAL   Dev. Biol. 239 (2), 241-256 (2001)
  MEDLINE   21643879
   PUBMED   11784032
COMMENT     Contact: Masanori Taira
            Department of Biological Sciences
            Graduate School of Science, University of Tokyo; CREST, Japan
            Science and Technology Corporation, Japan
            7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
            Tel: 81-03-5841-4434
            Fax: 81-03-5841-4434
            Email: m.taira@biol.s.u-tokyo.ac.jp,
            URL:http://www.biol.s.u-tokyo.ac.jp/users/lmb/lmb-hp.html.
FEATURES             Location/Qualifiers
     source          1..359
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="p5el2"
                     /tissue_type="anterior endomesoderm"
                     /dev_stage="late gastrula to early neurula stages 12.5-13"
                     /clone_lib="Shibata Xenopus AEM lambda-ZAP II cDNA
                     library"

BASE COUNT        90 a        92 c        82 g        93 t        2 others
ORIGIN

Query Match         40.3%;   Score 26.6;  DB 9;    Length 359;
Best Local Similarity  63.1%;  Pred. No. 1.2e+02;
Matches    41;  Conservative   0;  Mismatches  24;  Indels   0;  Gaps   0;

Qy      2  TGAAAAAGACACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGG  61
           |||| ||| || ||||  ||||| || ||| || ||||  |||||| |||| ||||||||
Db    129  TGAAGAAGGCAGATATCGGTATCGCTATGGGCATCGCTGGTTCTGATGTCTCCAAGCAGG  188

Qy     62  CGGCC  66
           || |||
Db    189  CAGCC  193


Search completed: August 19, 2003, 08:29:19
Job time : 194.082 secs
```

OM protein - protein search, using sw model

```
          Copyright  GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.
```

Run on: August 8, 2003, 17:08:17 ; Search time 44 Seconds
                                    (without alignments)
                                    1359.998 Million cell updates/sec

Title: US-09-987-455-8
Perfect score: 2061
Sequence: 1 MKKTAIAIAVALAGFATVAQ.........GVYTKVTNYLDWIRDNMRPG 377

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq 19Jun03:*
```
 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2061	100.0	377	23	AAE25034	OmpA-K2S fusion pr
2	1961	95.1	372	20	AAW87770	Human tissue plasm
3	1961	95.1	396	13	AAR25190	mt-PA6. Escherich
4	1961	95.1	438	10	AAP94409	Sequence of coding
5	1961	95.1	439	16	AAR68851	Delta 2-89 tissue
6	1961	95.1	446	11	AAR08150	Non-glycosylated t
7	1961	95.1	472	10	AAP94416	Sequence of coding
8	1961	95.1	519	16	AAR68852	Delta 55-62 tissue
9	1961	95.1	527	11	AAR05489	tPA024 precursor p
10	1961	95.1	527	11	AAR05488	tPA024 precursor p
11	1961	95.1	527	12	AAR13910	T-PA deriv. (I).
12	1961	95.1	527	12	AAR13911	T-PA deriv. (II).
13	1961	95.1	527	12	AAR13912	T-PA deriv. (III).
14	1961	95.1	527	12	AAR13913	T-PA deriv. (IV).
15	1961	95.1	527	12	AAR13914	T-PA deriv. (V).
16	1961	95.1	527	13	AAR20223	t-PA analogue expr
17	1961	95.1	527	13	AAR21593	tPA variant - E94A
18	1961	95.1	527	13	AAR21594	tPA variant - D95A
19	1961	95.1	527	13	AAR21595	tPA variant - D95G
20	1961	95.1	527	13	AAR21596	tPA variant - E94A
21	1961	95.1	527	13	AAR21600	tPA variant - E94A
22	1961	95.1	527	13	AAR22617	Mutated recombinan
23	1961	95.1	527	13	AAR22618	Mutated recombinan
24	1961	95.1	527	13	AAR22619	Mutated recombinan
25	1961	95.1	527	13	AAR22620	Mutated recombinan
26	1961	95.1	527	13	AAR22621	Mutated recombinan
27	1961	95.1	527	19	AAW45907	Single chain form
28	1961	95.1	527	23	AAE25044	Human tissue plasm
29	1961	95.1	527	23	AAE24190	Human tissue plasm
30	1961	95.1	527	23	AAG79362	Human tissue plasm
31	1961	95.1	528	16	AAR68853	Bifibronectin doma
32	1961	95.1	531	7	AAP60700	Sequence of tissue
33	1961	95.1	537	12	AAR11661	Tissue plasminogen
34	1961	95.1	539	12	AAR12846	T-PA growth domain
35	1961	95.1	552	12	AAR11662	Tissue plasminogen
36	1961	95.1	556	12	AAR13148	T-PA variant contg
37	1961	95.1	557	12	AAR13020	T-PA variant contg
38	1961	95.1	557	12	AAR13149	T-PA variant contg
39	1961	95.1	558	12	AAR13153	T-PA with -ve char
40	1961	95.1	559	12	AAR12340	T-PA variant contg
41	1961	95.1	559	12	AAR13151	T-PA with -ve char
42	1961	95.1	560	12	AAR12341	T-PA variant contg
43	1961	95.1	561	12	AAR12367	T-PA variant contg
44	1961	95.1	562	7	AAP60790	Sequence of human
45	1961	95.1	562	7	AAP60214	Sequence of active

ALIGNMENTS

```
RESULT 1
AAE25034
ID   AAE25034 standard; Protein; 377 AA.
XX
AC   AAE25034;
XX
DT   30-OCT-2002  (first entry)
XX
DE   OmpA-K2S fusion protein.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; fusion protein.
XX
OS   Unidentified.
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   N-PSDB; AAD40614.
XX
```

PT　Producing active, correctly folded recombinant tissue plasminogen
PT　activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT　the protein-encoding DNA operably linked to DNA coding for signal
PT　peptide.OmpA -
XX
PS　Claim 25; Page 35-36; 80pp; English.
XX
CC　The present invention relates to a method of producing extracellularly
CC　secreted, active, correctly folded, recombinant tissue plasminogen
CC　activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC　variants in prokaryotic cells by expressing the protein-encoding DNA
CC　operably linked to DNA coding for signal peptide OmpA. The method is
CC　useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC　Sequences of the invention are useful for manufacturing a medicament
CC　for treating stroke, cardiac infarction, acute myocardial infarction,
CC　pulmonary embolism, any artery occlusion such as intracranial artery
CC　occlusion (e.g., arteries supplying the brain), peripherally occluded
CC　arteries, coronary artery occlusion, deep vein thrombosis or related
CC　diseases associated with unwanted blood clotting. The present sequence
CC　is a fusion protein comprising OmpA and K2S protein.
XX
SQ　Sequence 377 AA;

```
Query Match             100.0%;  Score 2061;  DB 23;  Length 377;
Best Local Similarity   100.0%;  Pred. No. 1.2e-167;
Matches   377;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 MKKTAIAIAVALAGFATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIG  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKKTAIAIAVALAGFATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIG  60

Qy   61 KYVTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KYVTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQ 120

Qy  121 PQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 PQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPH 180

Qy  181 HLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVV 240

Qy  241 RTVCLPPADLQLPDWTECELSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTV 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 RTVCLPPADLQLPDWTECELSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTV 300

Qy  301 TDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVY 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVY 360

Qy  361 TKVTNYLDWIRDNMRPG 377
          |||||||||||||||||
Db  361 TKVTNYLDWIRDNMRPG 377
```

RESULT 2
AAW87770
ID　AAW87770 standard; Protein; 372 AA.
XX
AC　AAW87770;
XX
DT　29-MAR-1999 (first entry)
XX
DE　Human tissue plasminogen activator.
XX
KW　Tissue plasminogen activator-like protease; t-PALP; human;
KW　circulatory system-related disorder; blood clotting; stroke;
KW　thrombosis; peripheral arterial occlusion; pulmonary embolism;
KW　myocardiothrombosis; diagnosis; therapy.
XX
OS　Homo sapiens.
XX

PN　WO9854199-A1.
XX
PD　03-DEC-1998.
XX
PF　27-MAY-1998;　98WO-US10728.
XX
PR　28-MAY-1997;　97US-0048000.
XX
PA　(HUMA-) HUMAN GENOME SCI INC.
XX
PI　Ebner R, Moore PA, Ruben SM;
XX
DR　WPI; 1999-070207/06.
XX
PT　New tissue plasminogen activator-like protease - useful in the
PT　diagnosis and treatment of circulatory system-related disorders
XX
PS　Disclosure; Page 57-58; 76pp; English.
XX
CC　This is the amino acid sequence of the translation product of the
CC　human mRNA for tissue plasminogen activator (tPA). It shares
CC　homology with a novel human polypeptide (see AAW87769), designated
CC　tissue plasminogen activator-like protease (t-PALP). The homology
CC　between t-PALP and tPA indicates that t-PALP may be involved in
CC　the regulation of normal and abnormal clotting in e.g. stroke,
CC　deep-vein thrombosis, peripheral arterial occlusion, pulmonary
CC　embolism and myocardiothrombosis.
XX
SQ　Sequence 372 AA;

```
Query Match             95.1%;  Score 1961;  DB 20;  Length 372;
Best Local Similarity   99.7%;  Pred. No. 4.2e-159;
Matches   355;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY  80
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY  76

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 136

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 196

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 256

Qy  261 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 316

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 372
```

RESULT 3
AAR25190
ID　AAR25190 standard; Protein; 396 AA.
XX
AC　AAR25190;
XX
DT　25-MAR-2003 (updated)
DT　11-DEC-1992 (first entry)
XX
DE　mt-PA6.
XX
KW　trp; operon; tPA; finger; Kringle domain; EGF; plasminogen; fibrinogen.
XX
OS　Escherichia coli.
XX

XX
PN   EP493037-A2.
XX
PD   01-JUL-1992.
XX
PF   20-DEC-1991;   91EP-0311894.
XX
PR   24-DEC-1990;   90US-0633584.
XX
PA   (ELIL) LILLY & CO ELI.
XX
PI   Burck PJ, Van Jackson C, Smith GF;
XX
DR   WPI; 1992-219148/27.
XX
PT   New di:glycosylated tissue plasminogen activator derivs. - for
PT   thromboembolic treatment e.g. deep vein thrombosis, acute
PT   myocardial infarction and thrombotic stroke
XX
PS   Claim 2; Page 31; 48pp; English.
XX
CC   The sequence given is a modified tissue plasminogen activator (tPA),
CC   mt-PA6. mt-PA6 is diglycosylated, and lacks the finger, EGF and
CC   Kringle 1 domains. It may be used as a thrombolytic agent. The novel
CC   tPA posseses the properties of sparing plasminogen and fibrinogen
CC   while displaying a longer in vivo half life and providing greater
CC   maintenance of coronary blood flow.
CC   (Updated on 25-MAR-2003 to correct PN field.)
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 396 AA;

Query Match   95.1%; Score 1961; DB 13; Length 396;
Best Local Similarity 99.7%; Pred. No.4.6e-159;
Matches 355; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
              ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   41 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 100

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 160

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 220

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 280

Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   281 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 340

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   341 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 396

RESULT 4
AAP94409
ID   AAP94409 standard; protein; 438 AA.
XX
AC   AAP94409;
XX
DT   25-MAR-2003   (updated)
DT   03-OCT-2002   (updated)
DT   18-JUN-1990   (first entry)
XX
DE   Sequence of coding region in plasmid pTQkPA delta trp.
XX

KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   N-PSDB; AAN91123.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page -; 68pp; English.
XX
CC   Coding region of plasmid pTQkPA delta trp.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 438 AA;

Query Match   95.1%; Score 1961; DB 10; Length 438;
Best Local Similarity 99.7%; Pred. No. 5.1e-159;
Matches 355; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
              ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   82 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 141

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 201

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 261

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 321

Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 381

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 437

RESULT 5
AAR68851
ID   AAR68851 standard; protein; 439 AA.
XX
AC   AAR68851;
XX
DT   25-MAR-2003   (updated)
DT   22-NOV-1995   (first entry)
XX
DE   Delta 2-89 tissue plasminogen activator.

```
XX
KW   Human; des-epidermal growth factor homologous plasminogen activator;
KW   tPA; liver membrane; reduced affinity; EGF homologous; thrombosis;
KW   thrombolytic; increased half-life.
XX
OS   Homo sapiens (engineered).
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 1..2
FT                  /note= "amino acids 2-89 of tPA have been deleted"
XX
PN   US5376547-A.
XX
PD   27-DEC-1994.
XX
PF   29-JAN-1988;    88US-0150267.
XX
PR   30-JAN-1987;    87US-0008795.
PR   29-JAN-1988;    88US-0150267.
XX
PA   (AMHP ) AMERICAN HOME PROD CORP.
XX
PI   Hung PP, Kalyan NK, Lee SL;
XX
DR   WPI; 1995-043464/06.
XX
PT   New modified plasminogen activator cpds. - having regions removed
PT   to reduce affinity for liver membranes and increase circulation
PT   half-life.
XX
PS   Claim 1; Page ?; 26pp; English.
XX
CC   Amino acid residues 2-89 contain the fibronectin and EGF regions of
CC   human tPA. Deletion of these regions results in a tissue plasminogen
CC   activator with reduced affinity for liver cell membranes; the
CC   mutant protein is not cleared from the circulation as rapidly as is
CC   wild-type tPA. The specification only gives the sequence around the
CC   deletion and not the full-length sequence of "delta 2-89 tPA"; the
CC   sequence in AAR68851 has been obtained by amending a previously
CC   disclosed wild-type human tPA sequence (from DE3930099) according to
CC   the description given in Example 1.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 439 AA;

Query Match       95.1%;  Score 1961; DB 16; Length 439;
Best Local Similarity 99.7%;  Pred. No. 5.2e-159;
Matches 355; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy    21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    84 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 143

Qy    81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   144 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 203

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   204 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 263

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 323

Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   324 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 383

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   384 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 439
```

```
RESULT 6
AAR08150
ID   AAR08150 standard; protein; 446 AA.
XX
AC   AAR08150;
XX
DT   09-JAN-2003  (updated)
DT   01-MAR-1991  (first entry)
XX
DE   Non-glycosylated tPA deriv. lacking finger- and EGF-domains.
XX
KW   Tissue plasminogen activator; thrombolysis; finger domain;
KW   EGF domain.
XX
OS   Synthetic.
XX
PN   EP400545-A.
XX
PD   05-DEC-1990.
XX
PF   28-MAY-1990;    90EP-0110096.
XX
PR   14-JUL-1989;    89DE-3923339.
PR   31-MAY-1989;    89DE-3917781.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Stern A, Kohnert U, Rudolph R, Fischer S, Martin U;
XX
DR   WPI; 1990-363094/49.
DR   N-PSDB; AAQ06762.
XX
PT   New non-glycosylated form of tissue plasminogen activator - with
PT   thrombolytic activity and long plasma life
XX
PS   Claim 1; page 15; 21pp; German.
XX
CC   This non-glycosylated tPA deriv. lacks the finger- and EGF-domains.
CC   It has a lower clearance rate (longer half-life) than natural tPA
CC   while retaining thrombolytic activity and stimulation by fibrin.
CC   It is produced by recombinant methods and site-specific muta-
CC   genesis.
CC   (Updated on 09-JAN-2003 to add missing OS field.)
XX
SQ   Sequence 446 AA;

Query Match       95.1%;  Score 1961; DB 11; Length 446;
Best Local Similarity 99.7%;  Pred. No. 5.3e-159;
Matches 355; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy    21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    91 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 150

Qy    81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   151 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 210

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 270

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   271 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 330

Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   331 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 390

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   391 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 446
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 7
AAP94416
ID   AAP94416 standard; protein; 472 AA.
XX
AC   AAP94416;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pmTQk112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   N-PSDB; AAN91133.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page -; 68pp; English.
XX
CC   Coding region of plasmid pmTQk112.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 472 AA;

Query Match   95.1%;   Score 1961; DB 10; Length 472;
Best Local Similarity   99.7%;   Pred. No. 5.6e-159;
Matches 355;   Conservative   0;   Mismatches 1;   Indels   0;   Gaps   0;

```
Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  117 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 176

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  177 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 236

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVPGEEQK 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  237 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVPGEEEQK 296

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  297 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 356

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  357 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 416
```

```
Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  417 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 472
```

RESULT 8
AAR68852
ID   AAR68852 standard; protein; 519 AA.
XX
AC   AAR68852;
XX
DT   25-MAR-2003 (updated)
DT   22-NOV-1995 (first entry)
XX
DE   Delta 55-62 tissue plasminogen activator.
XX
KW   Human; des-epidermal growth factor homologous plasminogen activator;
KW   tPA; liver membrane; reduced affinity; EGF homologous; thrombosis;
KW   thrombolytic; increased half-life.
XX
OS   Homo sapiens (engineered).
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 54..55
FT   /note= "amino acids 55-62 of tPA have been deleted"
XX
PN   US5376547-A.
XX
PD   27-DEC-1994.
XX
PF   29-JAN-1988;   88US-0150267.
XX
PR   30-JAN-1987;   87US-0008795.
PR   29-JAN-1988;   88US-0150267.
XX
PA   (AMHP) AMERICAN HOME PROD CORP.
XX
PI   Hung PP, Kalyan NK, Lee SL;
XX
DR   WPI; 1995-043464/06.
XX
PT   New modified plasminogen activator cpds. - having regions removed
PT   to reduce affinity for liver membranes and increase circulation
PT   half-life.
XX
PS   Claim 1; Page ?; 26pp; English.
XX
CC   Amino acid residues 55-62 fall within the EGF region of human tPA.
CC   Deletion of these residues results in a tissue plasminogen
CC   activator with reduced affinity for liver cell membranes; the
CC   mutant protein is not cleared from the circulation as rapidly as is
CC   wild-type tPA. The specification only gives the sequence around the
CC   deletion and not the full-length sequence of "delta 55-62 tPA"; the
CC   sequence in AAR68852 has been obtained by amending a previously
CC   disclosed wild-type human tPA sequence (from DE3930099) according
CC   to the description given in Example 6.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 519 AA;

Query Match   95.1%;   Score 1961;   Length 519;
Best Local Similarity   99.7%;   Pred. No. 6.3e-159; DB 16;
Matches 355;   Conservative   0;   Mismatches 1;   Indels   0;   Gaps   0;

```
Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  164 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 223

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  224 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 283

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVPGEEBQK 200
```

```
Db  172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNY 231
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGLFADIASHPWQA 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGLFADIASHPWQA 291
Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPFPHHLTVILGRTYRVVPGEEEQK 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPFPHHLTVILGRTYRVVPGEEEQK 351
Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411
Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471
Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527
```

```
RESULT 10
AAR05488
ID   AAR05488 standard; protein; 527 AA.
XX
AC   AAR05488;
XX
DT   25-MAR-2003 (updated)
DT   18-OCT-1990 (first entry)
XX
DE   tPA024 precursor protein.
XX
KW   Tissue plasminogen activator; TPA;fibrin;
KW   myocardial infarction; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   67
FT                   /note="Asn linked to oligosaccharide"
XX
PN   EP373896-A.
XX
PD   20-JUN-1990.
XX
PF   12-DEC-1989;    89EP-0312993.
XX
PR   12-DEC-1988;    88JP-0314172.
XX
PA   (YAMA ) YAMANOUCHI PHARM CO LTD.
PA   (YAWA ) NIPPON STEEL CORP.
XX
PI   Shimizu Y, Yano E, Yano S, Kato M, Kinoshita A, Kawasaki T;
PI   Ishida J, Gushima H;
XX
DR   WPI; 1990-187265/25.
DR   N-PSDB; AAQ04904.
XX
PT   New tissue plasminogen activator analogues -
PT   contg. glycosylated asparagine unit for increased half-life.
XX
PS   Disclosure; Page ?; 34pp; English.
XX
CC   The tPA analogue carries an oligosaccharide side chain linked to an
CC   Asn residue replacing either Tyr-67 or Ile-86. The nature of the
CC   side chain depends on the host cell used.
CC   The analogue has a longer half life and may be administered at
CC   lower doses.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   527 AA;
```

```
Db      ||||||||||||||||||||||||||||||||||||||||||||||||
    284 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPFPHHLTVILGRTYRVVPGEEEQK 343
Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 403
Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  404 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 463
Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  464 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 519
```

```
RESULT 9
AAR05489
ID   AAR05489 standard; protein; 527 AA.
XX
AC   AAR05489;
XX
DT   25-MAR-2003 (updated)
DT   18-OCT-1990 (first entry)
XX
DE   tPA024 precursor protein.
XX
KW   Tissue plasminogen activator; TPA;fibrin;
KW   myocardial infarction; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   86
FT                   /note="Asn linked to oligosaccharide "
XX
PN   EP373896-A.
XX
PD   20-JUN-1990.
XX
PF   12-DEC-1989;    89EP-0312993.
XX
PR   12-DEC-1988;    88JP-0314172.
XX
PA   (YAMA ) YAMANOUCHI PHARM CO LTD.
PA   (YAWA ) NIPPON STEEL CORP.
XX
PI   Shimizu Y, Yano E, Yano S, Kato M, Kinoshita A, Kawasaki T;
PI   Ishida J, Gushima H;
XX
DR   WPI; 1990-187265/25.
DR   N-PSDB; AAQ04903.
XX
PT   New tissue plasminogen activator analogues -
PT   contg. glycosylated asparagine unit for increased half-life.
XX
PS   Disclosure; Page ?; 34pp; English.
XX
CC   The tPA analogue carries an oligosaccharide side chain linked to an
CC   Asn residue replacing either Tyr-67 or Ile-86. The nature of the
CC   side chain depends on the host cell used.
CC   The analogue has a longer half life and may be administered at
CC   lower doses.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   527 AA;

Query Match             95.1%;  Score 1961;  DB 11;  Length 527;
Best Local Similarity   99.7%;  Pred. No. 6.4e-159;
Matches   355;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNY 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Query Match   95.1%; Score 1961; DB 11; Length 527;
Best Local Similarity 99.7%; Pred. No. 6.4e-159;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 231

Qy    81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 351

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy   261 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 376
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 527
```

RESULT 11
AAR13910
ID   AAR13910 standard; protein; 527 AA.
XX
AC   AAR13910;
XX
DT   25-MAR-2003 (updated)
DT   25-NOV-1991 (first entry)
XX
DE   T-PA deriv. (I).
XX
KW   Tissue plasminogen activator; thrombolysis; myocardial infarction.
XX
OS   Homo sapiens.
XX
PN   EP445464-A.
XX
PD   11-SEP-1991.
XX
PF   15-NOV-1990;   90EP-0312445.
XX
PR   15-NOV-1990;   90EP-0312445.
PR   01-MAR-1990;   90JP-0050428.
PR   02-AUG-1990;   90JP-0206458.
XX
PA   (KANF) KANEGAFUCHI KAGAKU KOGYO KK.
XX
PI   Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
PI   Kakutani T;
XX
DR   WPI; 1991-268391/37.
XX
PT   New tissue plasminogen derivs. - having amino acid replacements
PT   to increase persistence in the blood while retaining thrombolytic
PT   activity
XX
PS   Claim 4; Page 29; 39pp; English.
XX
CC   The t-PA has the following amino acids of the mature protein
CC   replaced: N(37) for S, S(38) for V, G(39) for V, R(40) for E,
CC   A(41) for S and Q(42) for S. The deriv. has improved persistence in
CC   the blood compared with native t-PA, while retaining the same level of
CC   thrombolytic activity. It can be used for the treatment of
CC   thrombotic disorders such as myocardial infarction.
CC   The wild-type t-PA sequence was retrieved by the indexer
CC   (DE3930099) and "mutated" according to the claims.
CC   See also AAR13910-14.
XX   (Updated on 25-MAR-2003 to correct PA field.)
SQ   Sequence 527 AA;

Query Match   95.1%; Score 1961; DB 12; Length 527;
Best Local Similarity 99.7%; Pred. No. 6.4e-159;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 231

Qy    81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 351

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy   261 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 376
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 527
```

RESULT 12
AAR13911
ID   AAR13911 standard; Protein; 527 AA.
XX
AC   AAR13911;
XX
DT   25-MAR-2003 (updated)
DT   25-NOV-1991 (first entry)
XX
DE   T-PA deriv. (II).
XX
KW   Tissue plasminogen activator; thrombolysis; myocardial infarction.
XX
OS   Homo sapiens.
XX
PN   EP445464-A.
XX
PD   11-SEP-1991.
XX
PF   15-NOV-1990;   90EP-0312445.
XX
PR   15-NOV-1990;   90EP-0312445.
PR   01-MAR-1990;   90JP-0050428.
PR   02-AUG-1990;   90JP-0206458.
XX
PA   (KANF) KANEGAFUCHI KAGAKU KOGYO KK.
XX
PI   Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
PI   Kakutani T;
XX
DR   WPI; 1991-268391/37.
XX
PT   New tissue plasminogen derivs. - having amino acid replacements
PT   to increase persistence in the blood while retaining thrombolytic
PT   activity
XX

Left column

```
PS   Claim 5; Page 29; 39pp; English.
XX
CC   The t-PA has the following amino acids of the mature protein
CC   replaced: N(37) for S, S(38) for V, G(39) for E,
CC   A(41) for F and Q(42) for S.  The deriv. has improved persistence in
CC   the blood compared with native t-PA, while retaining the same level of
CC   thrombolytic activity.  It can be used for the treatment of
CC   thrombotic disorders such as myocardial infarction.
CC   The wild-type t-PA sequence was retrieved by the indexer
CC   (DE3930099) and "mutated" according to the claims.
CC   See also AAR13910-14.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   527 AA;

Query Match             95.1%;  Score 1961;  DB 12;  Length 527;
Best Local Similarity   99.7%;  Pred. No. 6.4e-159;
Matches  355;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps    0;

Qy   21  AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY  80
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172  ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231

Qy   81  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQRIKGGLFADIASHPWQA 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  232  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQRIKGGLFADIASHPWQA 291

Qy  141  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  292  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 351

Qy  201  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  352  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy  261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  412  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy  321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  472  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527

RESULT 13
AAR13912
ID   AAR13912 standard; Protein; 527 AA.
XX
AC   AAR13912;
XX
DT   25-MAR-2003 (updated)
DT   25-NOV-1991 (first entry)
XX
DE   T-PA deriv. (III).
XX
KW   Tissue plasminogen activator; thrombolysis; myocardial infarction.
XX
OS   Homo sapiens.
XX
PN   EP445464-A.
XX
PD   11-SEP-1991.
XX
PF   15-NOV-1990;    90EP-0312445.
XX
PR   15-NOV-1990;    90EP-0312445.
PR   01-MAR-1990;    90JP-0050428.
PR   02-AUG-1990;    90JP-0206458.
XX
PA   (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX
PI   Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
PT   Kakutani T;
```

Right column

```
XX
DR   WPI; 1991-268391/37.
XX
PT   New tissue plasminogen derivs. - having amino acid replacements
PT   to increase persistence in the blood while retaining thrombolytic
PT   activity
XX
PS   Claim 6; Page 29; 39pp; English.
XX
CC   The t-PA has the following amino acids of the mature protein
CC   replaced: N(37) for S, S(38) for V, R(40) for I,
CC   A(41) for V and Q(42) for S.  The deriv. has improved persistence in
CC   the blood compared with native t-PA, while retaining the same level of
CC   thrombolytic activity.  It can be used for the treatment of
CC   thrombotic disorders such as myocardial infarction.
CC   The wild-type t-PA sequence was retrieved by the indexer
CC   (DE3930099) and "mutated" according to the claims.
CC   See also AAR13910-14.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   527 AA;

Query Match             95.1%;  Score 1961;  DB 12;  Length 527;
Best Local Similarity   99.7%;  Pred. No. 6.4e-159;
Matches  355;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps    0;

Qy   21  AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY  80
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172  ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231

Qy   81  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQRIKGGLFADIASHPWQA 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  232  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQRIKGGLFADIASHPWQA 291

Qy  141  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  292  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 351

Qy  201  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  352  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy  261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  412  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy  321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  472  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527

RESULT 14
AAR13913
ID   AAR13913 standard; Protein; 527 AA.
XX
AC   AAR13913;
XX
DT   25-MAR-2003 (updated)
DT   25-NOV-1991 (first entry)
XX
DE   T-PA deriv. (IV).
XX
KW   Tissue plasminogen activator; thrombolysis; myocardial infarction.
XX
OS   Homo sapiens.
XX
PN   EP445464-A.
XX
PD   11-SEP-1991.
XX
PF   15-NOV-1990;    90EP-0312445.
XX
PR   15-NOV-1990;    90EP-0312445.
```

```
PR  01-MAR-1990;  90JP-0050428.
PR  02-AUG-1990;  90JP-0206458.
XX
PA  (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX
PI  Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
PI  Kakutani T;
XX
DR  WPI; 1991-268391/37.
XX
PT  New tissue plasminogen derivs. - having amino acid replacements
PT  to increase persistence in the blood while retaining thrombolytic
PT  activity
XX
PS  Claim 7; Page 29; 39pp; English.
XX
CC  The t-PA has the following amino acids of the mature protein
CC  replaced: G(161) for R, K(162) for R and S(165) for W.
CC  The deriv. has improved persistence in the blood compared with native
CC  t-PA, while retaining the same level of thrombolytic activity.
CC  It can be used for the treatment of thrombotic disorders such as
CC  myocardial infarction.
CC  The wild-type t-PA sequence was retrieved by the indexer
CC  (DE3930099) and "mutated" according to the claims.
CC  See also AAR13910-14.
CC  (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ  Sequence  527 AA;

Query Match   95.1%;   Score 1961;   DB 12;   Length 527;
Best Local Similarity   99.7%;   Pred. No. 6.4e-159;
Matches   355;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 231

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 351

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527

RESULT 15
AAR13914
ID  AAR13914 standard; Protein; 527 AA.
XX
AC  AAR13914;
XX
DT  25-MAR-2003  (updated)
DT  25-NOV-1991  (first entry)
XX
DE  T-PA deriv. (V).
XX
KW  Tissue plasminogen activator; thrombolysis; myocardial infarction.
XX
OS  Homo sapiens.
XX
```

```
FH  Key               Location/Qualifiers
FT  Misc-difference   37
FT                    /label= ASN, SER
FT  Misc-difference   38
FT                    /label= SER, VAL
FT  Misc-difference   39
FT                    /label= GLY, VAL
FT  Misc-difference   40
FT                    /label= ARG, GLU
FT  Misc-difference   41
FT                    /label= ALA, SER
FT  Misc-difference   42
FT                    /label= GLN, SER
XX
PN  EP445464-A.
XX
PD  11-SEP-1991.
XX
PF  15-NOV-1990;  90EP-0312445.
XX
PR  15-NOV-1990;  90EP-0312445.
PR  01-MAR-1990;  90JP-0050428.
PR  02-AUG-1990;  90JP-0206458.
XX
PA  (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX
PI  Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
PI  Kakutani T;
XX
DR  WPI; 1991-268391/37.
XX
PT  New tissue plasminogen derivs. - having amino acid replacements
PT  to increase persistence in the blood while retaining thrombolytic
PT  activity
XX
PS  Claim 10; Page 29; 39pp; English.
XX
CC  The t-PA has the following amino acids of the mature protein
CC  replaced: N(115) for P and one of N(37) for S, S(38) for V, G(39)
CC  for V, R(40) for E, A(41) for S and Q(42) for S.
CC  The deriv. has improved persistence in the blood compared with native
CC  t-PA, while retaining the same level of thrombolytic activity.
CC  It can be used for the treatment of thrombotic disorders such as
CC  myocardial infarction.
CC  The wild-type t-PA sequence was retrieved by the indexer
CC  (DE3930099) and "mutated" according to the claims.
CC  See also AAR13910-14.
CC  (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ  Sequence  527 AA;

Query Match   95.1%;   Score 1961;   DB 12;   Length 527;
Best Local Similarity   99.7%;   Pred. No. 6.4e-159;
Matches   355;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 231

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 351

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy    321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCGQKDVPGVYTKVTNYLDWIRDNMRP 527


Search completed: August  8, 2003, 17:13:51
Job time : 45 secs
```

OM protein - protein search, using sw model

```
              GenCore version 5.1.6
   Copyright (c) 1993 - 2003  Compugen Ltd.
```

Run on: August 8, 2003, 17:13:07 ; Search time 20 Seconds
                            (without alignments)
                   797.560 Million cell updates/sec

Title: US-09-987-455-8
Perfect score: 2061
Sequence: 1 MKKTAIAIAVALAGFATVAQ.........GVVTKVTNYLDWIRDNMRPG 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Issued_Patents AA:*
  1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
  2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
  3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
  4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
  5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
  6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1961	95.1	472	2	US-08-811-949-63	Sequence 63, Appl
2	1961	95.1	527	1	US-07-609-510B-16	Sequence 16, Appl
3	1961	95.1	527	5	PCT-US91-01025A-2	Sequence 2, Appli
4	1961	95.1	527	6	5185259-8	Patent No. 5185259
5	1961	95.1	562	2	US-08-811-949-43	Sequence 43, Appl
6	1961	95.1	562	2	US-08-560-098A-50	Sequence 50, Appl
7	1961	95.1	562	2	US-08-883-795A-38	Sequence 38, Appl
8	1961	95.1	562	6	5185259-3	Patent No. 5185259
9	1961	95.1	562	6	5200340-2	Patent No. 5200340
10	1961	95.1	562	6	5344773-2	Patent No. 5344773
11	1958	95.0	562	6	5244676-5	Patent No. 5244676
12	1957	95.0	355	2	US-08-811-949-45	Sequence 45, Appl
13	1957	95.0	437	2	US-08-811-949-49	Sequence 49, Appl
14	1955.5	94.9	389	2	US-08-811-949-65	Sequence 65, Appl
15	1954	94.8	437	2	US-08-811-949-55	Sequence 55, Appl
16	1953	94.8	437	2	US-08-811-949-51	Sequence 51, Appl
17	1953	94.8	527	2	US-08-811-949-39	Sequence 39, Appl
18	1951.5	94.7	378	4	US-09-553-498-10	Sequence 10, Appl
19	1951.5	94.7	378	4	US-09-618-869-10	Sequence 10, Appl
20	1950	94.6	355	1	US-08-137-116-1	Sequence 1, Appli
21	1950	94.6	355	1	US-08-217-618-1	Sequence 1, Appli
22	1950	94.6	355	1	US-08-427-640-2	Sequence 2, Appli
23	1950	94.6	355	1	US-08-217-617A-1	Sequence 1, Appli
24	1950	94.6	355	1	US-08-217-616-1	Sequence 1, Appli
25	1950	94.6	355	2	US-08-811-949-53	Sequence 53, Appl
26	1950	94.6	355	3	US-08-794-528-1	Sequence 1, Appli
27	1950	94.6	355	6	5223256-1	Patent No. 5223256
28	1949	94.6	355	2	US-08-811-949-47	Sequence 47, Appl
29	1946	94.4	437	2	US-08-811-949-57	Sequence 57, Appl
30	1943.5	94.3	389	2	US-08-811-949-67	Sequence 67, Appl
31	1935	93.9	355	1	US-08-427-640-6	Sequence 6, Appli
32	1933.5	93.8	356	2	US-08-427-640-4	Sequence 4, Appli
33	1931	93.7	355	2	US-08-811-949-59	Sequence 59, Appl
34	1929	93.6	527	6	5520913-1	Patent No. 5520913
35	1919.5	93.1	354	2	US-08-811-949-61	Sequence 61, Appl
36	1882.5	91.3	347	2	US-08-811-949-1	Sequence 1, Appli
37	1787	86.7	356	1	US-08-427-640-8	Sequence 8, Appli
38	1704	82.7	326	4	US-09-411-977-3	Sequence 3, Appli
39	1458	70.7	383	3	US-08-558-269-6	Sequence 6, Appli
40	1458	70.7	383	3	US-09-410-882-6	Sequence 6, Appli
41	1423	69.0	477	2	US-08-560-098A-51	Sequence 51, Appl
42	1380	67.0	253	2	US-09-027-337-8	Sequence 8, Appli
43	1380	67.0	253	4	US-09-644-600-8	Sequence 8, Appli
44	1375	66.7	252	4	US-08-944-483-72	Sequence 72, Appl
45	880.5	42.7	546	6	5200340-6	Patent No. 5200340

ALIGNMENTS

RESULT 1
US-08-811-949-63
; Sequence 63, Application US/0811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
;   INFORMATION FOR SEQ ID NO: 63:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 472 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-811-949-63

Query Match   95.1%;   Score 1961; DB 2; Length 472;
Best Local Similarity   99.7%;   Pred. No. 8e-186;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTVTAQNPSAQALGLGKHNY 80

Db   117 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 176
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   177 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 236
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 296
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   297 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 356
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   357 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 416
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 472

RESULT 2
US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
;   APPLICANT: Berg et al.
;   TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: IN.
;   COUNTRY: U.S.A.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
;   COMPUTER: Macintosh
;   OPERATING SYSTEM: Macintosh
;   SOFTWARE: Microsoft Word
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/609,510B
;   FILING DATE: 19901106
;   CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 527 amino acids
;   TYPE: AMINO ACID
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-07-609-510B-16

Query Match   95.1%; Score 1961; DB 1; Length 527;
Best Local Similarity   99.7%; Pred. No. 9.3e-186;
Matches   355; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260

Db   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLLDWIRDNMRP 527

RESULT 3
PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
;   APPLICANT: Genentech, Inc.
;   TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
;   TITLE OF INVENTION: Specific Properties
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US91/01025A
;   FILING DATE: 19910214
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/486,657
;   FILING DATE: 1 March 1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Hasak, Janet E.
;   REGISTRATION NUMBER: 28,616
;   REFERENCE/DOCKET NUMBER: 454P2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/266-1896
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 527 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
PCT-US91-01025A-2

Query Match   95.1%; Score 1961; DB 5; Length 527;
Best Local Similarity   99.7%; Pred. No. 9.3e-186;
Matches   355; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   352  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471
Qy   321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527
```

RESULT 4
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.;KOHR, WILLIAM J.;PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
;   NUMBER OF SEQUENCES: 15
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/489,855
;   FILING DATE: 02-MAR-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 12,694
;   FILING DATE: 09-FEB-1987
;   APPLICATION NUMBER: 483,052
;   FILING DATE: 07-APR-1983
;   APPLICATION NUMBER: 398,003
;   FILING DATE: 14-JUL-1982
;   APPLICATION NUMBER: 374,860
;   FILING DATE: 05-MAY-1982
; SEQ ID NO:8:
;   LENGTH: 527
5185259-8

Query Match   95.1%; Score 1961; DB 6; Length 527;
Best Local Similarity 99.7%; Pred. No. 9.3e-186;
Matches 355; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231
Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 351
Qy 201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411
Qy 261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471
Qy 321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527
```

RESULT 5
US-08-811-949-43
; Sequence 43, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;     ADDRESSEE: P.C.
;     STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;     CITY: ARLINGTON
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/811,949
;     FILING DATE: 05-MAR-1997
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: OBLON, NORMAN F.
;     REGISTRATION NUMBER: 24,618
;     REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-413-3000
;     TELEFAX: 703-413-2220
;   INFORMATION FOR SEQ ID NO: 43:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 562 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-811-949-43

Query Match      95.1%;  Score 1961;  DB 2;  Length 562;
Best Local Similarity  99.7%;  Pred. No. 1e-185;
Matches  355;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy   21  AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207  ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 266
Qy   81  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326
Qy  141  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386
Qy  201  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446
Qy  261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506
Qy  321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

RESULT 6
US-08-560-098A-50
; Sequence 50, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
;   APPLICANT: WNENDT, Stephan
;   APPLICANT: HEINZEL-WIELAND, Regina
;   APPLICANT: STEFFENS, Gerd Josef

```
;  TITLE OF INVENTION:  Recombinant DNA Molecules and Expression
;  TITLE OF INVENTION:  Vectors for Tissue Plasminogen Activator
;  NUMBER OF SEQUENCES:  39
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  BERESKIN & PARR
;  STREET:  40 King Street West
;  CITY:  Toronto
;  STATE:  Ontario
;  COUNTRY:  Canada
;  ZIP:  M5H 3Y2
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  PatentIn Release #1.0,  Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/883,795A
;  FILING DATE:  27-JUN-1997
;  CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  Gravelle, Micheline
;  REGISTRATION NUMBER:  40,261
;  REFERENCE/DOCKET NUMBER:  7841-062
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  (416) 364-7311
;  TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  38:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  562 amino acids
;  TYPE:  amino acid
;  STRANDEDNESS:  single
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  protein (tPA)
US-08-883-795A-38

Query Match            95.1%;  Score 1961;  DB 2;  Length 562;
Best Local Similarity   99.7%;  Pred. No. 1e-185;
Matches   355;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY  80
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 266

Qy     81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy    141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386

Qy    201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy    261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy    321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562

RESULT 8
5185259-3
;Patent No. 5185259
;  APPLICANT:  GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
;VEHAR, GORDON A.
;  TITLE OF INVENTION:  TRUNCATED HUMAN TISSUE PLASMINOGEN
;ACTIVATOR
;  NUMBER OF SEQUENCES:  15
;  CURRENT APPLICATION DATA:
```

```
;  TITLE OF INVENTION:  Proteins having Fibrinolytic and
;  TITLE OF INVENTION:  Coagulation-inhibiting Properties
;  NUMBER OF SEQUENCES:  60
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Evenson, McKeown, Edwards & Lenahan
;  STREET:  1200 G Street, N.W., Suite 700
;  CITY:  Washington
;  STATE:  DC
;  COUNTRY:  USA
;  ZIP:  20005
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  PatentIn Release #1.0,  Version #1.30 (EPO)
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/560,098A
;  FILING DATE:  17-NOV-1995
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER:  P 44 40 892.7
;  FILING DATE:  17-NOV-1994
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  EVANS, Joseph D.
;  REGISTRATION NUMBER:  26,269
;  REFERENCE/DOCKET NUMBER:  148/42448
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  (202) 628-8800
;  TELEFAX:  (202) 628-8844
;  INFORMATION FOR SEQ ID NO:  50:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  562 amino acids
;  TYPE:  amino acid
;  STRANDEDNESS:  single
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  protein
US-08-560-098A-50

Query Match            95.1%;  Score 1961;  DB 2;  Length 562;
Best Local Similarity   99.7%;  Pred. No. 1e-185;
Matches   355;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY  80
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 266

Qy     81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy    141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386

Qy    201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy    261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy    321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562

RESULT 7
US-08-883-795A-38 ;
; Sequence 38, Application US/08883795A
; Patent No. 5985607
;  GENERAL INFORMATION:
;  APPLICANT:  Delcuve, Genevieve
;  APPLICANT:  Awang, Gregor
```

```
;           APPLICATION NUMBER: US/07/489,855
;           FILING DATE: 02-MAR-1990
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 12,694
;           FILING DATE: 09-FEB-1987
;           APPLICATION NUMBER: 483,052
;           FILING DATE: 07-APR-1983
;           APPLICATION NUMBER: 398,003
;           FILING DATE: 14-JUL-1982
;           APPLICATION NUMBER: 374,860
;           FILING DATE: 05-MAY-1982
;SEQ ID NO:3:
;  LENGTH: 562
5185259-3

Query Match        95.1%;    Score 1961; DB 6;  Length 562;
Best Local Similarity    99.7%;  Pred. No. 1e-185;
Matches 355;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
        | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

RESULT 9
5200340-2
```
;Patent No. 5200340
;   APPLICANT: FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
;   TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;   NUMBER OF SEQUENCES: 34
;   CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/07/53,412
;           FILING DATE: 22-MAY-1987
;SEQ ID NO:2:
;  LENGTH: 562
5200340-2

Query Match        95.1%;    Score 1961; DB 6;  Length 562;
Best Local Similarity    99.7%;  Pred. No. 1e-185;
Matches 355;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
        | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386
```

```
Db  327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

RESULT 10
5344773-2
```
;Patent No. 5344773
;   APPLICANT: WEI, CHA-MER;HSIUNG, NANCY;REDDY, VERMURI B.;
;LEMONTT, JEFFREY F.;DACKOWSKI, WILLIAM;DOUGLAS, RICHARD;
;COLE, EDWARD S.;PURCELL JR., RICHARD D.;LAU, DAVID TAI-YUI
;   TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
;ACTIVATOR PRODUCED BY RECOMBIANT DNA
;   NUMBER OF SEQUENCES: 6
;   CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/07/782,686
;           FILING DATE: 01-OCT-1985
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 656,770
;           FILING DATE: 01-OCT-1984
;SEQ ID NO:2:
;  LENGTH: 562
5344773-2

Query Match        95.1%;    Score 1961; DB 6;  Length 562;
Best Local Similarity    99.7%;  Pred. No. 1e-185;
Matches 355;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
        | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 386

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

RESULT 11
524676-5
```
;Patent No. 5244676
;   APPLICANT: BELL, LESLIE D.;MAYER, ERNEST J.;PALMIER, MARK O.
;TOLUNAY, H.ESER;WARREN, THOMAS G.;WUN, TZE-CHEIN
;   TITLE OF INVENTION: MODIFIED TISSUE PLASMINOGEN ACTIVATOR
;WITH MODIFIED GLYCOSYLATION SITE
;   NUMBER OF SEQUENCES: 9
;   CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/07/203,047
```

; FILING DATE: 06-JUN-1988
;SEQ ID NO:5:
; LENGTH: 562
5244676-5

```
Query Match          95.0%;   Score 1958;  DB 6;  Length 562;
Best Local Similarity 99.4%;  Pred. No. 2e-185;
Matches  354;  Conservative  1;  Mismatches  1;  Indels  0;  Gaps  0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
        : |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 SCSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 386

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

RESULT 12
US-08-811-949-45
; Sequence 45, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-45

```
Query Match          95.0%;   Score 1957;  DB 2;  Length 355;
Best Local Similarity 100.0%;  Pred. No. 1.3e-185;
Matches  354;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 82
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 61

Qy   83 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 142
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 121

Qy  143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 202
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 181

Qy  203 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 262
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 241

Qy  263 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 322
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 301

Qy  323 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 355
```

RESULT 13
US-08-811-949-49
; Sequence 49, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 49:

```
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 437 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-811-949-49

Query Match             95.0%;  Score 1957;  DB 2;  Length 437;
Best Local Similarity   99.4%;  Pred. No. 1.8e-185;
Matches  354;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 141

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  142 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 201

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  202 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTYRVVPGEEEQK 261

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 321

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  322 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 381

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  382 DACQGDSGGPLVCLNDGRITLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 437
```

RESULT 14
US-08-811-949-65

```
; Sequence 65, Application US/08811949
; Patent No. 5840533
;   GENERAL INFORMATION:
;     APPLICANT: NIWA, MINEO
;     APPLICANT: SAITO, YOSHIMASA
;     APPLICANT: SASAKI, HITOSHI
;     APPLICANT: HAYASHI, MASAKO
;     APPLICANT: NOTANI, JOUJI
;     APPLICANT: KOBAYASHI, MASAKAZU
;     TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;     NUMBER OF SEQUENCES: 67
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;        ADDRESSEE: P.C.
;        STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;        CITY: ARLINGTON
;        STATE: VA
;        COUNTRY: USA
;        ZIP: 22202
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/811,949
;        FILING DATE: 05-MAR-1997
;        CLASSIFICATION: 435
;     ATTORNEY/AGENT INFORMATION:
;        NAME: OBLON, NORMAN F.
;        REGISTRATION NUMBER: 24,618
;        REFERENCE/DOCKET NUMBER: 18-966-0
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: 703-413-3000
;        TELEFAX: 703-413-2220
```

```
; INFORMATION FOR SEQ ID NO: 65:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 389 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-811-949-65

Query Match             94.9%;  Score 1955.5;  DB 2;  Length 389;
Best Local Similarity   93.2%;  Pred. No. 2.1e-185;
Matches  359;  Conservative    4;  Mismatches   11;  Indels   11;  Gaps    2;

Qy    3 KTAIAIAVALAG--FATVAQ-------AASEGNSDCYFGNGSAYRGTHSLTESGASCL 51
         | :. : :|.| : |: :         |||||||||||||||||||||||||||||||
Db    5 KRGLCCVLLCGALFVSPSQEIHARFRRGARSEGNSDCYFGNGSAYRGTHSLTESGASCL 64

Qy   52 PWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS 111
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65 PWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS 124

Qy  112 TCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAH 171
         |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  125 TCGLRQYSQPQFRIKGGLFADIASHPWQAPIFAKHRRSPGERFLCGGILISSCWILSAAH 184

Qy  172 CFQERFPPHHLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSS 231
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  185 CFQERFPPHHLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSS 244

Qy  232 RCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCT 291
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  245 RCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCT 304

Qy  292 SQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGC 351
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  305 SQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGC 364

Qy  352 GQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||
Db  365 GQKDVPGVYTKVTNYLDWIRDNMRP 389
```

RESULT 15
US-08-811-949-55

```
; Sequence 55, Application US/08811949
; Patent No. 5840533
;   GENERAL INFORMATION:
;     APPLICANT: NIWA, MINEO
;     APPLICANT: SAITO, YOSHIMASA
;     APPLICANT: SASAKI, HITOSHI
;     APPLICANT: HAYASHI, MASAKO
;     APPLICANT: NOTANI, JOUJI
;     APPLICANT: KOBAYASHI, MASAKAZU
;     TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;     NUMBER OF SEQUENCES: 67
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;        ADDRESSEE: P.C.
;        STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;        CITY: ARLINGTON
;        STATE: VA
;        COUNTRY: USA
;        ZIP: 22202
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/811,949
;        FILING DATE: 05-MAR-1997
;        CLASSIFICATION: 435
;     ATTORNEY/AGENT INFORMATION:
;        NAME: OBLON, NORMAN F.
```

```
;     REGISTRATION NUMBER: 24,618
;     REFERENCE/DOCKET NUMBER: 18-966-0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 703-413-3000
;       TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 55:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 437 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-811-949-55
```

```
Query Match     94.8%;  Score 1954;  DB 2;  Length 437;
Best Local Similarity  99.4%;  Pred. No. 3.5e-185;
Matches  354;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 141

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  142 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFDIKGGLFADIASHPWQA 201

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEBEEQK 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  202 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEBEEQK 261

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 321

Qy  261 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  322 SGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 381

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  382 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 437
```

Search completed: August 8, 2003, 17:15:51
Job time : 21 secs

```
                    GenCore version 5.1.6

          Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:       August   8, 2003, 17:15:07 ; Search time 53 Seconds
                                  (without alignments)
                          844.763 Million cell updates/sec

Title:        US-09-987-455-8
Perfect score: 2061
Sequence:      1 MKKTAIAIAVALAGFATVAQ.........GVVTKVTNYLDWIRDNMRPG 377

Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5

Searched:     451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters:        451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published Applications AA:*
              1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
              2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
              3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
              4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
              5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
              6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
              7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
              8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
              9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
             10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
             11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
             12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
             13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
             14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
             15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
             16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
             17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
             18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2061	100.0	377	11	US-09-987-455-8	Sequence 8, Appli
2	1961	95.1	372	9	US-09-084-491A-3	Sequence 3, Appli
3	1961	95.1	372	14	US-10-102-704-3	Sequence 3, Appli
4	1961	95.1	527	11	US-09-987-457-18	Sequence 18, Appl
5	1961	95.1	527	11	US-09-987-457-19	Sequence 19, Appl
6	1961	95.1	562	9	US-09-969-271-7	Sequence 7, Appli
7	1961	95.1	562	15	US-10-193-656-8	Sequence 8, Appli
8	1957	95.0	354	11	US-09-987-457-10	Sequence 10, Appl
9	1957	95.0	354	11	US-09-987-455-11	Sequence 11, Appl
10	1952	94.7	562	10	US-09-974-298-145	Sequence 145, App
11	1875.5	91.0	343	11	US-09-987-457-14	Sequence 14, Appl
12	1875.5	91.0	343	11	US-09-987-455-15	Sequence 15, Appl
13	1865.5	90.5	343	11	US-09-987-457-15	Sequence 15, Appl
14	1865.5	90.5	343	11	US-09-987-455-16	Sequence 16, Appl
15	1850.5	89.8	339	11	US-09-987-457-12	Sequence 12, Appl
16	1850.5	89.8	339	11	US-09-987-455-13	Sequence 13, Appl
17	1837	89.1	335	11	US-09-987-457-13	Sequence 13, Appl
18	1837	89.1	335	11	US-09-987-455-14	Sequence 14, Appl
19	1829	88.7	331	11	US-09-987-457-11	Sequence 11, Appl
20	1829	88.7	331	11	US-09-987-455-12	Sequence 12, Appl
21	1704	82.7	308	11	US-09-987-457-16	Sequence 16, Appl
22	1704	82.7	308	11	US-09-987-455-17	Sequence 17, Appl
23	1704	82.7	326	14	US-10-057-951-3	Sequence 3, Appli
24	1465	71.1	268	11	US-09-987-457-17	Sequence 17, Appl
25	1465	71.1	268	11	US-09-987-455-18	Sequence 18, Appl
26	1341	65.1	246	11	US-09-898-837A-46	Sequence 46, Appl
27	784	38.0	403	10	US-09-880-503-6	Sequence 6, Appli
28	780	37.8	411	10	US-09-880-503-3	Sequence 3, Appli
29	780	37.8	431	14	US-10-076-421-2	Sequence 2, Appli
30	780	37.8	431	15	US-10-171-311-184	Sequence 184, App
31	777	37.7	431	15	US-10-193-656-4	Sequence 4, Appli
32	768	37.3	431	10	US-09-264-468B-1	Sequence 1, Appli
33	738	35.8	655	15	US-09-912-559-3	Sequence 28, Appl
34	709.5	34.4	560	10	US-09-912-559-3	Sequence 3, Appli
35	709.5	34.4	560	15	US-10-172-712-32	Sequence 32, Appl
36	698.5	33.9	560	10	US-09-912-559-4	Sequence 4, Appli
37	671.5	32.6	615	11	US-09-858-909-2	Sequence 2, Appli
38	671.5	32.6	615	15	US-10-172-712-30	Sequence 30, Appl
39	600	29.1	812	9	US-09-788-142-1	Sequence 1, Appli
40	600	29.1	812	9	US-09-761-120-1	Sequence 1, Appl
41	600	29.1	812	9	US-09-873-676-81	Sequence 81, Appl
42	600	29.1	812	15	US-09-335-325-1	Sequence 1, Appli
43	600	29.1	812	15	US-10-131-241-1	Sequence 1, Appli
44	593	28.8	323	10	US-09-880-503-7	Sequence 7, Appli
45	583	28.3	276	10	US-09-880-503-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-09-987-455-8
; Sequence 8, Application US/09987455
; Publication No. US2003004972A1
; GENERAL INFORMATION:
;  APPLICANT: Aranya Manosroi
;  APPLICANT: Jiradej Manosroi
;  APPLICANT: Chatchai Tayapiwatana
;  APPLICANT: Friedrich Goetz
;  APPLICANT: Rolf-Guenther Werner
;  TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;  TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;  FILE REFERENCE: 0652.2190001
;  CURRENT APPLICATION NUMBER: US/09/987,455
;  CURRENT FILING DATE: 2001-11-14
;  PRIOR APPLICATION NUMBER: 60/268,574
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: GB 0027779.8
;  PRIOR FILING DATE: 2000-11-14
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;  LENGTH: 377
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: OmpA-K2S
;  OTHER INFORMATION: fusion protein
US-09-987-455-8

  Query Match             100.0%;  Score 2061;  DB 11;  Length 377;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-193;
  Matches  377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MKKTAIAIAVALAGFATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MKKTAIAIAVALAGFATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIG 60
```

```
Qy   61 KVTTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KVTTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQ 120

Qy  121 PQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 PQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPH 180

Qy  181 HLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 HLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVV 240

Qy  241 RTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTV 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 RTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTV 300

Qy  301 TDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVY 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVY 360

Qy  361 TKVTNYLDWIRDNMRPG 377
          |||||||||||||||||
Db  361 TKVTNYLDWIRDNMRPG 377
```

RESULT 2
US-09-084-491A-3
; Sequence 3, Application US/09084491A
; Patent No. US2002061576A1
; GENERAL INFORMATION:
;   APPLICANT: MOORE, PAUL A.
;   APPLICANT: RUBEN, STEVEN M.
;   APPLICANT: EBNER, REINHARD
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: HUMAN GENOME SCIENCES, INC.
;   STREET: 9410 KEY WEST AVENUE
;   CITY: ROCKVILLE
;   STATE: MD
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/084,491A
;   FILING DATE: 27-MAY-1998
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: BROOKES, ANDERS A.
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PF378
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 372 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-084-491A-3

Query Match   95.1%; Score 1961; DB 9; Length 372;
Best Local Similarity   99.7%; Pred. No. 4.9e-183;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 76

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 136

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 196

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 256

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 316

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 372
```

RESULT 3
US-10-102-704-3
; Sequence 3, Application US/10102704
; Publication No. US20020164768A1
; GENERAL INFORMATION:
;   APPLICANT: Moore et al.
;   TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
;   FILE REFERENCE: PF378C1
;   CURRENT APPLICATION NUMBER: US/10/102,704
;   CURRENT FILING DATE: 2002-03-22
;   PRIOR APPLICATION NUMBER: 09/084,491
;   PRIOR FILING DATE: 1998-05-27
;   PRIOR APPLICATION NUMBER: 60/048,000
;   PRIOR FILING DATE: 1997-05-28
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 372
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-102-704-3

Query Match   95.1%; Score 1961; DB 14; Length 372;
Best Local Similarity   99.7%; Pred. No. 4.9e-183;
Matches 355; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 76

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 136

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 196

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 256

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 316

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 372
```

RESULT 4
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens (tPA)
US-09-987-457-18

Query Match   95.1%;   Score 1961; DB 11; Length 527;
Best Local Similarity 99.7%;   Pred. No. 7.9e-183;
Matches 355; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHHLTVILGRTYRVVPGEEEQK 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHHLTVILGRTYRVVPGEEEQK 351

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527

RESULT 5
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15

PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match   95.1%;   Score 1961; DB 11; Length 527;
Best Local Similarity 99.7%;   Pred. No. 7.9e-183;
Matches 355; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 231

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 291

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHHLTVILGRTYRVVPGEEEQK 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHHLTVILGRTYRVVPGEEEQK 351

Qy   201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 411

Qy   261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471

Qy   321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 527

RESULT 6
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US2002009817A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match   95.1%;   Score 1961; DB 9; Length 562;
Best Local Similarity 99.7%;   Pred. No. 8.6e-183;
Matches 355; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy   141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHHLTVILGRTYRVVPGEEEQK 200

```
Db  327  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 386
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  201  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

```
RESULT 7
US-10-193-656-8
; Sequence 8, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
;  APPLICANT: NY, Tor
;  APPLICANT: HOLMDAHL, Rikard
;  APPLICANT: LI, Jinan
;  TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
;  FILE REFERENCE: 3810/IJ577-US3
;  CURRENT APPLICATION NUMBER: US/10/193,656
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: US 60/304,461
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/304,490
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/305,182
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/305,182
;  PRIOR FILING DATE: 2001-07-13
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 8
;  LENGTH: 562
;  TYPE: PRT
;  ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
;  DATABASE ACCESSION NUMBER: GenBank / P00750
;  DATABASE ENTRY DATE: 1986-07-21
;  RELEVANT RESIDUES: (1)..(562)
US-10-193-656-8

Query Match         95.1%;  Score 1961;  DB 15;  Length 562;
Best Local Similarity 99.7%;  Pred. No. 8.6e-183;
Matches  355;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;
```

```
Qy   21  AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db  207  ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 266

Qy   81  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy  141  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 386

Qy  201  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562
```

```
RESULT 8
US-09-987-457-10
; Sequence 10, Application US/09987457
; Publication No. US2003013150A1
; GENERAL INFORMATION:
;  APPLICANT: Manosroi, Aranya
;  APPLICANT: Manosroi, Jiradej
;  APPLICANT: Tayapiwatana, Chatchai
;  APPLICANT: Goetz, Friedrich
;  APPLICANT: Werner, Rolf-Guenther
;  TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
;  FILE REFERENCE: 0652.2180001
;  CURRENT APPLICATION NUMBER: US/09/987,457
;  CURRENT FILING DATE:  2001-11-14
;  PRIOR APPLICATION NUMBER: 60/268,573
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: GB 00 27 782.2
;  PRIOR FILING DATE: 2000-11-14
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 10
;  LENGTH: 354
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: part of the
;  OTHER INFORMATION: recombinant K2S molecule
US-09-987-457-10

Query Match         95.0%;  Score 1957;  DB 11;  Length 354;
Best Local Similarity 100.0%;  Pred. No. 1.1e-182;
Matches  354;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy   23  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 82
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 60

Qy   83  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 120

Qy  143  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFB 180

Qy  203  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 240

Qy  263  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 300

Qy  323  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 354
```

```
RESULT 9
US-09-987-455-11
; Sequence 11, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
;  APPLICANT: Aranya Manosroi
;  APPLICANT: Jiradej Manosroi
;  APPLICANT: Chatchai Tayapiwatana
;  APPLICANT: Friedrich Goetz
;  APPLICANT: Rolf-Guenther Werner
;  TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;  TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;  FILE REFERENCE: 0652.2190001
```

```
;  CURRENT APPLICATION NUMBER: US/09/987,455
;  CURRENT FILING DATE: 2001-11-14
;  PRIOR APPLICATION NUMBER: 60/268,574
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: GB 0027779.8
;  PRIOR FILING DATE: 2000-11-14
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 11
;  LENGTH: 354
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: K2S 174-527
US-09-987-455-11

Query Match             95.0%;  Score 1957;  DB 11;  Length 354;
Best Local Similarity   100.0%;  Pred. No. 1.1e-182;
Matches   354;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      23  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR   82
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR   60

Qy      83  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  142
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  120

Qy     143  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  202
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  180

Qy     203  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  262
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  240

Qy     263  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  322
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  300

Qy     323  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  376
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  354
```

```
RESULT 10
US-09-974-298-145
;  Sequence 145, Application US/09974298
;  Patent No. US20020156263A1
;  GENERAL INFORMATION:
;  APPLICANT: Chen, Huei-Mei
;  ORGANISM: Homo sapiens
;  TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
;  FILE REFERENCE: PA-0037 P
;  CURRENT APPLICATION NUMBER: US/09/974,298
;  CURRENT FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,331
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
;  SEQ ID NO 145
;  LENGTH: 562
;  TYPE: PRT
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Incyte ID No. US2002015623A1 1001470CD1
US-09-974-298-145

Query Match             94.7%;  Score 1952;  DB 10;  Length 562;
Best Local Similarity   99.4%;  Pred. No. 6.5e-182;
Matches   354;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy      21  AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY   80
```

```
Db     207  ACSEGNSDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVYTAQNPSAQALGLGKHNY  266
                 |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||

Qy      81  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA  140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267  CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA  326

Qy     141  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK  200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327  AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK  386

Qy     201  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL  260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     387  FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL  446

Qy     261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH  320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     447  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH  506

Qy     321  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  376
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     507  DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  562
```

```
RESULT 11
US-09-987-457-14
;  Sequence 14, Application US/09987457
;  Publication No. US2003001315OA1
;  GENERAL INFORMATION:
;  APPLICANT: Manosroi, Aranya
;  APPLICANT: Manosroi, Jiradej
;  APPLICANT: Tayapiwatana, Chatchai
;  APPLICANT: Goetz, Friedrich
;  APPLICANT: Werner, Rolf-Guenther
;  TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
;  FILE REFERENCE: 0652.2180001
;  CURRENT APPLICATION NUMBER: US/09/987,457
;  CURRENT FILING DATE: 2001-11-14
;  PRIOR APPLICATION NUMBER: 60/268,573
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: GB 00 27 782.2
;  PRIOR FILING DATE: 2000-11-14
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 14
;  LENGTH: 343
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:   part of the
;  OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-14

Query Match             91.0%;  Score 1875.5;  DB 11;  Length 343;
Best Local Similarity   96.9%;  Pred. No. 9.7e-175;
Matches   343;  Conservative   0;  Mismatches   0;  Indels   11;  Gaps   1;

Qy      23  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR   82
                ||||||          ||||||||||||||||||||||||||||||||||||||||||
Db       1  SEGNSD--------THSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR   49

Qy      83  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  142
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  109

Qy     143  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  202
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  169

Qy     203  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  262
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  229
```

```
Qy    263  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    230  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  289

Qy    323  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  376
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    290  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  343
```

RESULT 12
US-09-987-455-15
; Sequence 15, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: K2S 191-527,
;   OTHER INFORMATION: modified
US-09-987-455-15

```
Query Match             91.0%;  Score 1875.5;  DB 11;  Length 343;
Best Local Similarity   96.9%;  Pred. No. 9.7e-175;
Matches  343;  Conservative   0;  Mismatches   0;  Indels  11;  Gaps   1;

Qy     23  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNYCR  82
           ||||||                  ||||||||||||||||||||||||||||||||||||
Db      1  SEGNSD----------THSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNYCR  49

Qy     83  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     50  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  109

Qy    143  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    110  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  169

Qy    203  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    170  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  229

Qy    263  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    230  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  289

Qy    323  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  376
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    290  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  343
```

RESULT 13
US-09-987-457-15
; Sequence 15, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:

;   APPLICANT: Manosroi, Aranya
;   APPLICANT: Manosroi, Jiradej
;   APPLICANT: Tayapiwatana, Chatchai
;   APPLICANT: Goetz, Friedrich
;   APPLICANT: Werner, Rolf-Guenther
;   TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
;   FILE REFERENCE: 0652.2180001
;   CURRENT APPLICATION NUMBER: US/09/987,457
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,573
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 00 27 782.2
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 18
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: part of the
;   OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-15

```
Query Match             90.5%;  Score 1865.5;  DB 11;  Length 343;
Best Local Similarity   96.6%;  Pred. No. 9.2e-174;
Matches  342;  Conservative   0;  Mismatches   0;  Indels  11;  Gaps   1;

Qy     23  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNYCR  82
           ||||||                  ||||||||||||||||||||||||||||||||||||
Db      1  SEGNSD----------THSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNYCR  49

Qy     83  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     50  NPDGDAKPWCHVLKNRRLTWEYCDVPSSSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  109

Qy    143  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    110  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  169

Qy    203  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    170  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  229

Qy    263  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    230  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  289

Qy    323  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  376
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    290  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  343
```

RESULT 14
US-09-987-455-16
; Sequence 16, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 191-527,
; OTHER INFORMATION: modified
US-09-987-455-16

Query Match             90.5%;  Score 1865.5; DB 11;  Length 343;
Best Local Similarity   96.6%;  Pred. No. 9.2e-174;
Matches   342;  Conservative   0;  Mismatches  1;  Indels  11;  Gaps   1;

Qy   23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 82
           |||||                 ||||||||||||||||||||||||||||||||||||||
Db    1 SEGNSD-------------THSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 49

Qy   83 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 142
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   50 NPDGDAKPWCHVLKNRRLTWEYCDVPSSSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 109

Qy  143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 169

Qy  203 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  170 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 229

Qy  263 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  230 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 289

Qy  323 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 343
```

RESULT 15
US-09-987-457-12

```
; SEQ ID NO 12
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:  part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-12

Query Match             89.8%;  Score 1850.5; DB 11;  Length 339;
Best Local Similarity   95.8%;  Pred. No. 2.6e-172;
Matches   339;  Conservative   0;  Mismatches  0;  Indels  15;  Gaps   1;

Qy   23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 82
           |||||                 ||||||||||||||||||||||||||||||||||||||
Db    1 SEGN------------------SLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 45

Qy   83 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   46 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 105

Qy  143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  106 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE 165

Qy  203 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 225

Qy  263 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  226 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 285

Qy  323 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  286 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 339
```

Search completed: August 8, 2003, 17:23:42
Job time : 54 secs

THIS PAGE BLANK (USPTO)

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     August  8, 2003, 17:12:22 ; Search time 19 Seconds
                                     (without alignments)
                                  1908.189 Million cell updates/sec

Title:        US-09-987-455-8
Perfect score: 2061
Sequence:      1 MKKTAIAIAVALAGFATVAQ.........GVYTKVTNYLDWIRDNMRPG 377

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Database :     PIR.76:*
                   1: pir1:*
                   2: pir2:*
                   3: pir3:*
                   4: pir4:*

Searched:      283308 seqs, 9168682 residues

Total number of hits satisfying chosen parameters:      283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	% Match	Length	DB	ID	Description
1	1961	95.1		562	1	UKHUT	t-plasminogen acti
2	1638.5	79.5		559	1	A35029	t-plasminogen acti
3	1600	77.6		559	1	A29941	t-plasminogen acti
4	1451.5	70.4		394	2	JS0600	t-plasminogen acti
5	1428	69.3		477	2	JS0598	t-plasminogen acti
6	1427	69.2		431	1	JS0599	t-plasminogen acti
7	1423	69.0		477	1	A34369	t-plasminogen acti
8	1389.5	67.4		477	2	JS0597	t-plasminogen acti
9	779.5	37.8		433	1	UKMS	u-plasminogen acti
10	777	37.7		431	1	UKHU	u-plasminogen acti
11	774.5	37.6		433	1	UKBAY	u-plasminogen acti
12	772.5	37.5		432	1	S18932	u-plasminogen acti
13	769.5	37.3		442	1	UKPG	u-plasminogen acti
14	762	37.0		434	1	A35005	u-plasminogen acti
15	745	36.1		433	1	JN0560	u-plasminogen acti
16	738	35.8		655	1	A46688	hepatocyte growth
17	709.5	34.4		560	2	JC4795	plasma hyaluronan-
18	698.5	33.9		558	1	JC5878	plasma hyaluronan-
19	671.5	32.6		615	1	KFHUL2	coagulation factor
20	670	32.5		603	2	S28941	coagulation factor
21	620.5	30.1		593	2	S45281	coagulation factor
22	600	29.1		812	1	PLMS	plasmin (EC 3.4.21
23	580.5	28.2		460	2	B61545	plasmin (EC 3.4.21
24	563.5	27.3		810	2	B30848	plasmin (EC 3.4.21
25	561.5	27.2		810	1	PLHU	plasmin (EC 3.4.21
26	552.5	26.8		790	1	PLPG	plasmin (EC 3.4.21
27	550.5	26.7		812	1	PLBO	plasmin (EC 3.4.21
28	547	26.5		810	2	I46260	plasmin (EC 3.4.21
29	541.5	26.3		4548	1	S00657	apoprotein(a) (EC
30	540	26.2		455	2	A61545	plasmin (EC 3.4.21
31	519	25.2		1420	2	A32869	apolipoprotein(a)
32	465	22.6		761	2	JC5759	brain-specific ser
33	453	22.0		855	2	JC7731	membrane-bound arg
34	405	19.7		613	2	S15468	complement C3b/C4b
35	398.5	19.3		625	1	TBBO	thrombin (EC 3.4.2
36	396.5	19.2		618	2	A35827	thrombin (EC 3.4.2
37	395.5	19.2		1524	2	T30337	polyprotein - Afri
38	393.5	19.1		261	2	A25606	tissue kallikrein
39	393	19.1		622	1	TBHU	thrombin (EC 3.4.2
40	393	19.1		728	1	JH0579	hepatocyte growth
41	391	19.0		638	1	KQMSPL	plasma kallikrein
42	389	18.9		436	2	JX0172	acrosin (EC 3.4.21
43	389	18.9		728	1	A60185	hepatocyte growth
44	387	18.8		728	1	A35644	hepatocyte growth
45	386.5	18.8		625	1	KFHU1	coagulation factor

ALIGNMENTS

RESULT 1
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N;Alternate names: t-PA; tissue plasminogen activator
C;Species: Homo sapiens (man)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
C;Accession: A94004; A23529; JT0562; A93293; S02125; A91343; A93951; A91322; A54645; I60
R;Ny, T.; Elgh, F.; Lund, B.
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A;Reference number: A94004; MUID:84298137; PMID:6089198
A;Accession: A94004
A;Molecule type: DNA
A;Residues: 1-562 <NYT>
A;Cross-references: GB:L00141
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translati
R;Friezner Degen, S.J.; Rajput, B.; Reich, B.
J. Biol. Chem. 261, 6972-6985, 1986
A;Title: The human tissue plasminogen activator gene.
A;Reference number: A23529; MUID:86196143; PMID:3009482
A;Accession: A23529
A;Molecule type: DNA
A;Residues: 1-562 <DEG>
A;Cross-references: GB:K03021; NID:g339817; PIDN:AAA98809.1; PID:g339818
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A;Title: Purification and characterization of tissue plasminogen activator secreted by h
A;Reference number: JT0562; MUID:91291340; PMID:1368681
A;Accession: JT0562
A;Molecule type: mRNA
A;Residues: 31-562 <ITA>
A;Cross-references: DDBJ:D01096; NID:g220128; PIDN:BAA00881.1; PID:g441174
A;Experimental source: embryonic lung fibroblast IMR-90 cells
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A;Reference number: A93293; MUID:83115262; PMID:6337343
A;Accession: A93293
A;Molecule type: mRNA
A;Residues: 1-562 <PEN>
A;Cross-references: GB:L00141
A;Experimental source: melanoma cells
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A;Reference number: S02125; MUID:88262579; PMID:3133640
A;Accession: S02125
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-562 <SAS>
A;Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244

A;Experimental source: fetal lung cells
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
FBBS Lett. 189, 145-149, 1985
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A;Reference number: A91343; MUID:85285620; PMID:3896853
A;Accession: A91343
A;Molecule type: mRNA
A;Residues: 1-38,'G',86-433,'E',435-562 <KAG>
A;Experimental source: Detroit 562 cells; ATCC 138
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
A;Reference number: A93951; MUID:83169656; PMID:6572897
A;Accession: A93951
A;Molecule type: mRNA
A;Residues: 251-358 <EDL>
A;Experimental source: melanoma cells
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A;Reference number: A90488; MUID:85000468; PMID:6433976
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
A;Reference number: A91322; MUID:84158956; PMID:6538514
A;Accession: A91322
A;Molecule type: protein
A;Residues: 33-45;311-320 <POH>
A;Experimental source: uterus
A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A;Reference number: A37567; MUID:87033611; PMID:3021732
A;Contents: annotation; fibrin binding site
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
EMBO J. 5, 3525-3530, 1986
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
A;Reference number: A37568; MUID:87161761; PMID:3030730
A;Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type p
A;Reference number: A60902; MUID:89044681; PMID:3142086
A;Contents: annotation; novel forms of expressed recombinant t-PA
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.;
Mol. Biol. Med. 3, 279-292, 1986
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expr
A;Reference number: A54645; MUID:86284200; PMID:3090401
A;Accession: A54645
A;Molecule type: mRNA
A;Residues: 1-562 <HAR>
A;Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032
A;Note: parts of this sequencing were confirmed by peptide sequencing
R;Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
A;Reference number: I60110; MUID:88054470; PMID:2824147
A;Accession: I60110
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-562 <RES>
A;Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A;Title: Isolation and characterization of the human tissue-type plasminogen activator g
A;Reference number: I55232; MUID:85289338; PMID:3161893
A;Accession: I55232
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-36 <RE2>
A;Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839

C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.
C;Genetics:
A;Gene: GDB:PLAT
A;Cross-references: GDB:119496; OMIM:173370
A;Map position: 8p12-8p12
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510.
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homo
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-32/Domain: propeptide #status predicted <PRO>
F;33-562/Product: t-plasminogen activator #status experimental <MAT>
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F;41-78/Domain: fibronectin type I repeat homology <1F1>
F;86-119/Domain: EGF homology <EGF>
F;127-208/Domain: kringle homology <KR1>
F;215-296/Domain: kringle homology <KR2>
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F;311-556/Domain: trypsin homology <TRY>
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299-4
F;152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F;357,406/Active site: His, Asp #status predicted
F;513/Active site: Ser #status experimental

Query Match   95.1%;   Score 1961; DB 1; Length 562;
Best Local Similarity   99.7%;   Pred. No. 3.1e-162;
Matches 355; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   21 AASBGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILGKVYTAQNPSAQALGLGKHNY 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 207 ACSGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQOPFRIKGGLFADIASHPWQA 326

Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEQK 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEQK 386

Qy 201 FEVEKYIVHKEFDDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 387 FEVEKYIVHKEFDDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy 261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy 321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQXDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQXDVPGVYTKVTNYLDWIRDNMRP 562

RESULT 2
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
A;Accession: A35029; A31597
R;Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A;Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A;Reference number: A35029; MUID:90130448; PMID:2105315
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226
R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988

A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activato
A;Reference number: A31597; MUID:89170114; PMID:3148445
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379,'K',381-559 <NYT>
A;Cross-references: GB:M23697; NID:g530159; PIDN:AAA41812.1; PID:g530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match   79.5%;   Score 1638.5; DB 1; Length 559;
Best Local Similarity   82.5%;   Pred. No. 3.1e-134;
Matches 287; Conservative 29; Mismatches 31; Indels 1; Gaps 1;

Qy   28 DCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGD 87
              ||||||||||||||||||||||||||||||||||||| || : ||||||||||||||||
Db   212 DCYVGKGVTYRGTHSFTTSKASCLPWNSMILIGKTYTAWRANSQALGLGRHNYCRNPDGD 271

Qy   88 AKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAIPAKGR 147
             |||||| | ||||||||||||||||||||||||||||| ||||||:|||||||||||
Db   272 AKPWCHVMKDRKLTWEYCDMSPCSTCGLRQYKPQFRIKGGLFTDITSHPWQAAIFVKNK 331

Qy   148 RSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKPEVEKYI 207
             ||||||||||||||||||||||||||||| || :||||||||||||||||| :||||
Db   332 RSPGERFLCGGVLISSCWVLSAAHCFVERFPPHHLKVVLGRTYRVVPGEEEQTFEIEKYI 391

Qy   208 VHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHE 267
             ||||||||||||||||||||||||:|| |||:||||||||||||||||||||||||||
Db   392 VHKEFDDDTYDNDIALLQLQLRSDSSQCAQESSVGTACLPDPDVQLPDWTECELSGYGKHE 451

Qy   268 ALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQGDS 327
             |: |:|||| |||||||||||||||||||||||||||||||||| :||||||||||||
Db   452 ASSPFFSDRLKEAHVRLYPSSRCTSQHLFNKTITSNMLCAGDTRTGGNQ-DVHDACQGDS 510

Qy   328 GGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMR 375
             ||||||| || ||||||||:||||||||||||| |:||||||||::||:
Db   511 GGPLVCMIDKRMTLLGIISWGLGCGQKDVPGIYTKVTNYLNWIQDNMK 558

RESULT 3
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29941; S48205; S48207; S48206
R;Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR
A;Reference number: A29941; MUID:88087303; PMID:2826484
A;Accession: A29941
A;Molecule type: mRNA
A;Residues: 1-559 <RIC>
A;Cross-references: GB:J03520; NID:g202109; PIDN:AAA40470.1; PID:g202110
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076; PMID:7523120
A;Accession: S48205
A;Molecule type: protein
A;Residues: 33-37,'X',39-40 <LIJ>

A;Accession: S48207
A;Molecule type: protein
A;Residues: 309-316 <LI2>
A;Accession: S48206
A;Molecule type: protein
A;Residues: 33-37,'X',39-40 <LIW>
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match   77.6%;   Score 1600; DB 1; Length 559;
Best Local Similarity   79.5%;   Pred. No. 6.8e-131;
Matches 283; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

Qy   21 AASEGNS-DCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHN 79
              :||::: ||||||||||||||||||||||||||||||:||||||:||||||||||||
Db   204 ACPKGKSEDCYVGKGVTYRGTHSLTTSQASCLPWNSIVLMGKSYTAWRTNSQALGLARHN 263

Qy   80 YCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQ 139
           |||||||||||||||| | |||||||||||||||||||||| |||||||||::||||||
Db   264 YCRNPDGDARPWCHVMKDRKLTWEYCDMSPCSTCGLRQYKRPQFRIKGGLYTDITSHPWQ 323

Qy   140 AAIPAKGRRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQ 199
           |||::|:| ||||||||||||:||||||||||||| || :|||||||:|||||||||||||
Db   324 AAIFVKNKRSPGERFLCGGVLISSCWVLSAAHCFLERFPPNHLKVVLGRTYRVVPGEEEQ 383

Qy   200 KFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECE 259
             |:|||||||||||||||||||||||||| :|| || ||:|||| ||||||||||||
Db   384 TFEIEKYIVHEEFDDDTYDNDIALLQLRSQSKQCAQESSVGTACLPDPNLQLPDWTECE 443

Qy   260 LSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANL 319
           |||||||||: |:|||| |||||||||||||||||||| :||:||||||||||: |||||
Db   444 LSGYGKHEASSPFFSDRLKEAHVRLYPSSRCTSQHLFNKTVTNNMLCAGDTRSGGNQ-DL 502

Qy   320 HDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMR 375
           ||||||||||||||| || ||||||||:||||||||||||| |:||||||||::||:
Db   503 HDACQGDSGGPLVCMINKQMTLTGIISWGLGCGQKDVPGVYTKVTNYLNWIDHDNMK 558

RESULT 4
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0600
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0600
A;Molecule type: mRNA
A;Residues: 1-394 <KRA>
A;Cross-references: GB:M63990; NID:g166078; PIDN:AAA31595.1; PID:g166079
A;Note: the authors translated the codon ATC for residue 75 as Thr
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-394/Product: plasminogen activator gamma #status predicted <PLA>

```
F;45-126/Domain: kringle homology <KRG>
F;143-388/Domain: trypsin homology <TRY>
F;45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F;142-143/Cleavage site: His-Ser (plasmin) #status predicted
F;189,238,345/Active site: His, Asp, Ser #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match     70.4%;   Score 1451.5;  DB 2;  Length 394;
Best Local Similarity   67.6%;   Pred. No. 3.5e-118;
Matches  263;  Conservative  42;  Mismatches  69;  Indels  15;  Gaps  2;

Qy      3  KTAIAIAVALAG---------FATVAQAASEGNSDCYFGNGSAYRGTHSLTESGAS   49
              |: :|| ||  | :|  :|:|| | |||||||||||||| |||||||
Db      6  KTKLLCVLLCGAVFSLPRQETYRQLARGSRAYGDPHATCYKDQGVTYRGTWSTSESGAQ   65

Qy     50  CLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPS  109
              |: | || | :   :   |||||||| :|: | ||| | ||||||||||||
Db     66  CINWNSNLLIRRTYNGRMPEAVKLGLGNHNYCRNPDGASKPWCYVIKARKFTSESCSVPV  125

Qy    110  CS--TCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWIL  167
              ||   ||||:|||: ||||| |||| :||:||||||| ||||||||||||||||
Db    126  CSKATCGLRKYKEPQLHSTGGLFTDITSHPWQAIFAQNRRSSGERFLCGGILISSCWVL  185

Qy    168  SAAHCFQERFPPHLTVILGRTYRVVPGBEEQKFEVEKYIVHKEFDDDTYNDIALLQLK  227
              |:||||||||||||||||||||  |::||| ::||:||::||||||||||||||||
Db    186  TAAHCFQERYPPQHLRVVLGRTYRVKPGKEEQTFEVEKCIVHEEFDDDTYNNDIALLQLK  245

Qy    228  SDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPS  287
              |||:||||||  ||||||| |||||||||||||||:||||| |||||:|:||| ||||||
Db    246  SGSPQCAQESDSVRAICLPEANLQLPDWTECELSGYGKHKSSSPFYSEQLKEGHVRLJYPS  305

Qy    288  SRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISW  347
              |||::|||||| ||:||||||||||||  |||||||||||||||||:|:|:|||||||||
Db    306  SRCTSKFLPNKTVTNNMLCAGDTRSGEIYPNVHDACQGDSGGPLVCMNDNHMTLLGIISW  365

Qy    348  GLGCGQKDVPGVTKVTNYLDWIRDNMRP  376
              | ||| :| |||||||||| |||||
Db    366  GVGCGEKDIPGVTKVTNYLGWIRDNMRP  394


RESULT 5
JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992  #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0598
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0598
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075
C;Superfamily: tissue plasminogen activator; EGF homology; kringle; serine proteinase
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F;185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match     69.3%;   Score 1428;  DB 2;  Length 477;
Best Local Similarity   70.2%;   Pred. No. 4.8e-116;
Matches  254;  Conservative  40;  Mismatches  66;  Indels  2;  Gaps  1;
```

```
Qy     17  TVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLG   76
               : :.|| ||  ||   :|||  :|||  :     :|| |
Db    116  TGKQCEVDTHATCYKDQGVTYRGTWSTSESGAQCINWNSNLLTRRTYNGRRSDAITLGLG  175

Qy     77  KHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPQFRIKGGLFADIA  134
               |||||||||| :||||  |:| |||||  :|:   ||||:|||: |||||||| ||||
Db    176  NHNYCRNPDNNSKPWCYVIKASKPILEFCSVPVCSKATCGLRKYKEPQLHSTGGLFTDIT  235

Qy    135  SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP  194
               |||||||||| ||||||||||||||||||||:||||||||||  |||||||||||||||
Db    236  SHPWQAAIFAQNRRSSGERFLCGGILISSCWVLTAAHCFQERYPPQHLRVVLGRTYRVKP  295

Qy    195  GEEEQKFEVEKYIVHKEFDDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPD  254
               |::||| ::||:||::|||||||||||||||||:|||||| ||||||| |||||||
Db    296  GKEEQTFEVEKCIVHEEFDDDTYNNDIALLQLKSGSPQCAQESDSVRAICLPEANLQLPD  355

Qy    255  WTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGG  314
               |||||||||||  |||||:|:||| |||||||||||::|||||| ||:|||||||||
Db    356  WTECELSGYGKHKSSSPFYSEQLKEGHVRLYPSSRCTSKFLFNKTVTNNMLCAGDTRSGE  415

Qy    315  PQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRDNM  374
               |  |||||||||||||||:|:|:|||||||||| ||| :| ||||||||| ||||||
Db    416  IYPNVHDACQGDSGGPLVCMNDNHMTLLGIISWGVGCGEKDIPGVTKVTNYLGWIRDNM  475

Qy    375  RP  376
               ||
Db    476  RP  477


RESULT 6
JS0599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992  #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0599
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.;
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0599
A;Molecule type: mRNA
A;Residues: 1-431 <KRA>
A;Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077
C;Superfamily: tissue plasminogen activator; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-431/Product: plasminogen activator beta #status predicted <PLA>
F;41-74/Domain: EGF homology <EGF>
F;82-163/Domain: kringle homology <KRG>
F;180-425/Domain: trypsin homology <TRY>
F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon
F;139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;179-180/Cleavage site: His-Ser (plasmin) #status predicted
F;226,275,382/Active site: His, Asp, Ser #status predicted
F;345-361,378-406/Disulfide bonds: #status predicted

Query Match     69.2%;   Score 1427;  DB 2;  Length 431;
Best Local Similarity   69.9%;   Pred. No. 5.3e-116;
Matches  253;  Conservative  41;  Mismatches  66;  Indels  2;  Gaps  1;
```

```
Qy     17  TVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLG   76
               : :.|| ||  ||   :|||  :|||  :     :|| |
Db     70  TGKQCEVDTHATCYKDQGVTYRGTWSTSESGAQCINWNSNLLTRRTYNGRRSDAITLGLG  129

Qy     77  KHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPQFRIKGGLFADIA  134
               |||||||||| :||||  |:| |||||  :|:   ||||:|||: |||||||| ||||
Db    130  NHNYCRNPDNNSKPWCYVIKASKPILEFCSVPVCSKATCGLRKYKEPQLHSTGGLFTDIT  189

Qy    135  SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP  194
               |||||||||| ||||||||||||||||||||:||||||||||  |||||||||||||||
Db    190  SHPWQAAIFAQNRRSSGERFLCGGILISSCWVLTAAHCFQERYPPQHLRVVLGRTYRVKP  249
```

Qy   375 RP 376
         ||
Db   476 RP 477

RESULT 8
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0597
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0597
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F;153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match   67.4%;   Score 1389.5;   DB 2;   Length 477;
Best Local Similarity   69.2%;   Pred. No. 1.1e-112;
Matches   252;   Conservative   37;   Mismatches   72;   Indels   3;   Gaps   2;

Qy   13 AGFATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQA 72
             ||| : : :: ||||| |||||| |:|||||||||| |:|| |||:| ||
Db   113 AGY-TGKRCEVDTRATCYBCGQGVTYRGTWSTAESRVECINWNSSLLTRRTYNGRMPDAFN 171

Qy   73 LGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPQFRIKGGLF 130
            ||||:||||||||| | ||||||| | |:|| :: :| ||||||:|||||| | |
Db   172 LGLGNHNYCRNPNGAPKPWCYVIKAGKFTSESCSVPVCSKATCGLRKYKEPQLHSTGGLF 231

Qy   131 ADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTY 190
            || :|||||||||| |||| ||||||||||||||| | ||||||| || ||||||||
Db   232 TDITSHPWQAAIFAQNRRSSGERFLCGGILISSCWVLTAAHCFQESYLPDQLKVVLGRTY 291

Qy   191 RVVPGEEEOKFEVEKYIVHKEFPDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADL 250
            ||||||| || |||||||||||||:||||||||||||||| :| :|:|||||::|:
Db   292 RVKPGEEEQTFKVKKYIVHKEFPDDTYNNDIALLQLKSDPQCAQESDSVRAICLPEANL 351

Qy   251 QLPDWTECELSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDT 310
            |||||||||||||||||| |||||||||||||||||||||| |:|||||||||||||||
Db   352 QLPDWTECELSGYGKHKSSSPPFYSEQLKEGHVRLYPSSRCAPKFLFNKTVTNNMLCAGDT 411

Qy   311 RSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWI 370
            |||| |:|||||||||||||||| |||:|||||||||||||| :||||||||||||||
Db   412 RSGEIYPNVHDACQGDSGGPLVCMNDNHMTLLGIISWGVGCGEKDVPGVYTKVTNYLGWI 471

Qy   371 RDNM 374
            ||||
Db   472 RDNM 475

RESULT 9
UKMS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C;Accession: A29420; A24615

Qy   195 GEEEQKFEVEKYIVHKEFPDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPD 254
            |:||||||||||||||||||||:||||||||||||| ||:|| |:|||||||||||
Db   250 GKEEQTFEVEKCIIHEEFPDDTYNNDIALLQLKSGSPQCAQESDSVRAICLPEANLQLPD 309

Qy   255 WTECELSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGG 314
            |||||||||||| |||||::: |||| ::||||||||| :|||||:|||||||||||||
Db   310 WTECELSGYGKHKSSSPPFYSEQLKEGHVRLYPSSRCTSKFLFNKTVTNNMLCAGDTRSGE 369

Qy   315 PQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNM 374
            |||||||||||||||||||| |||:|||||||||||||: :||||||||||||||||||
Db   370 IYPNVHDACQGDSGGPLVCMNDNHMTLLGIISWGVGCGEKDIPGVYTKVTNYLGWIRDNM 429

Qy   375 RP 376
            ||
Db   430 RP 431

RESULT 7
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C;Species: Megaderma lyra
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A34369
R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
J. Biol. Chem. 264, 17947-17952, 1989
A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A;Reference number: A34369; MUID:90036867; PMID:2509450
A;Accession: A34369
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <GAR>
A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match   69.0%;   Score 1423;   DB 1;   Length 477;
Best Local Similarity   70.2%;   Pred. No. 1.3e-115;
Matches   254;   Conservative   38;   Mismatches   68;   Indels   2;   Gaps   1;

Qy   17 TVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQAIGLG 76
            ||||||| |::::|||| ||||||||||||||||||||||||||||||||:||||:|:||
Db   116 TGKQCEVDTHATCYKDQGVTYRGTWSTSESGAQCINWNSNLLTRRTYNGRRSDAITLGLG 175

Qy   77 KHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPQFRIKGGLFADIA 134
            :|||||||| ::|:|| | | ||||||| | || ||||||||||||||||||||||
Db   176 NHNYCRNPDNNSKPWCYVIKASKFILEFCSVPVCSKATCGLRKYKEPQLHSTGGLFTDIT 235

Qy   135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
            |||||||||| |||| |||||||||||||||| | |:||||| || | |:|||||| |
Db   236 SHPWQAAIFAQNRRSSGERFLCGGILISSCWVLTAAHCFQERYPPQHLRVVLGRTYRVKP 295

Qy   195 GEEEQKFEVEKYIVHKEFPDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPD 254
            |:||||||||||||||||||||:||||||||||||| ||:|| |:|||||||||||
Db   296 GKEEQTFEVEKCIVHEEFPDDTYNNDIALLQLKSGSPQCAQESDSVRAICLPEANLQLPD 355

Qy   255 WTECELSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGG 314
            |||||||||||| |||||::: |||| ::||||||||| :|||||:|||||||||||||
Db   356 WTECELSGYGKHKSSSPPFYSEQLKEGHVRLYPSSRCTSKFLFNKTVTNNMLCAGDTRSGE 415

Qy   315 PQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNM 374
            |||||||||||||||||||| |||:|||||||||||||: :||||||||||||||||||
Db   416 IHPNVHDACQGDSGGPLVCRNDNHMTLLGIISWGVGCGEKDIPGVYTKVTNYLGWIRDNM 475

R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A;Title: The murine urokinase-type plasminogen activator gene.
A;Reference number: A29420; MUID:88163489; PMID:2831940
A;Accession: A29420
A;Molecule type: DNA
A;Residues: 1-433 <DEG>
A;Cross-references: GB:M17922; NID:g202296; PIDN:AAA40539.1; PID:g202297
R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher,
Eur. J. Biochem. 148, 225-232, 1985
A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
A;Reference number: A24615; MUID:85179474; PMID:2985383
A;Accession: A24615
A;Molecule type: mRNA
A;Residues: 1-433 <BEL>
A;Cross-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128
C;Genetics:
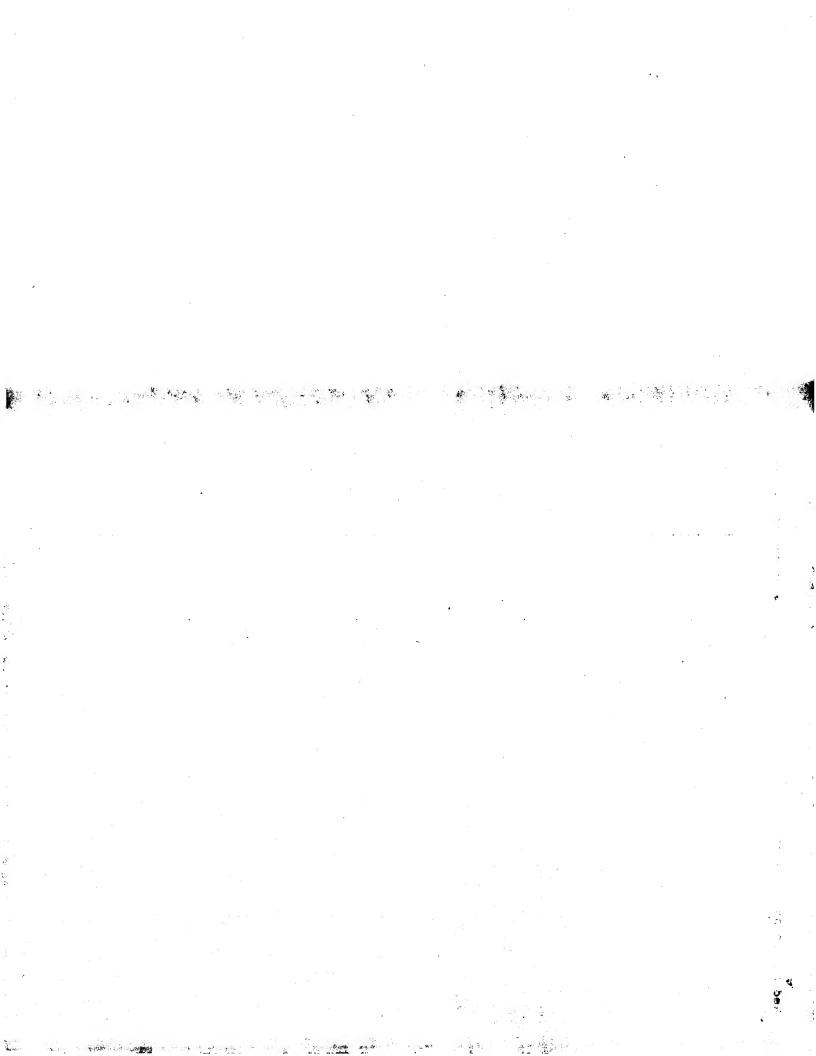A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;32-63/Domain: EGF homology <EGF>
F;71-152/Domain: kringle homology <KRG>
F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;180-421/Domain: trypsin homology <TRY>
F;169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

```
Query Match       37.8%;    Score 779.5; DB 1; Length 433;
Best Local Similarity  40.3%;   Pred. No. 8.6e-60;
Matches  148;   Conservative   67; Mismatches 129;   Indels  23; Gaps    5;

Qy      24 EGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRN 83
              :: || ||| :|||    ||  :: | | | :||||||:||||||
Db      66 DASKTCYHGNGDSYRGKANTDTKGRPCLAWNAPAVLQKPYNAHRPDAISLGLGKHNYCRN 125

Qy      84 PDGDAKPWCHVLKNRRLTWEYCDVPSCS----------------TCGLRQYSQPQFRIKGGL 129
              :||||             :||||  :: :||||||||||
Db     126 PDNQKRPWCYVQIGLRQFVQECMVHDCSLSKKPSSSVDQQGFQCG-QKALRPRFKIVGGE 184

Qy     130 FADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRT 189
              | ||:|| ||: ||     || || ||||||||    || |||| |||
Db     185 FTEVENQPWFAALYQKNKGGSPPSFKCGGSLISPCWVASAAHCFIQLPKKENYVVYLGQS 244

Qy     190 YRVVPGEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPP 247
              |||||       |:|: : ||||| :||||  :|: |||
Db     245 KESSYNPGEMKFEVEQLILHEYYREDSLAYHNDIALLKIRTSTGQCAQPSRSIQTICLPP 304

Qy     248 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 307
              |||||:  :: ||||:||:   : :||  |||||||||:||||
Db     305 RFTDAPFGSDCEITGFGKESESDYLYPKNLKMSVVKLVSHEQCMQPHYYGSEINYKMLCA 364

Qy     308 GDTRSGGPQANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYL 367
              || ||||||| ||||  :|||||||| |   ::|||  ||||||:|   |||||:|::|
Db     365 AD-----PEWKT-DSCKGDSGGPLICNIEGRPTLSGIVSWGRGCAEKNKPGVYTRVSHFL 418

Qy     368 DWIRDNM 374
              || :: :
Db     419 DWIQSHI 425
```

RESULT 10
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen act
in form
C;Species: Homo sapiens (man)
C;Date: 17-Dec-1982 #sequence_revision 04-Dec-1986 #text_change 15-Sep-2000
C;Accession: A00931; I52209; JT0102; A37561; I38102; S65783; A37562; A37563; A37564; A35
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.
A;Reference number: A00931; MUID:85215647; PMID:2987867
A;Accession: A00931
A;Molecule type: DNA
A;Residues: 1-431 <RIC>
A;Cross-references: GB:X02419; NID:g37601; PIDN:CAA26268.1; PID:g184524
A;Note: the authors translated the codon ATG for residue 214 as Ile
R;Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A;Reference number: I52209; MUID:86050639; PMID:3933505
A;Accession: I52209
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 145-161 <NAG1>
A;Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175
R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama,
Gene 36, 183-188, 1985
A;Title: Molecular cloning of cDNA coding for human preprourokinase.
A;Reference number: JT0102; MUID:86056954; PMID:2415429
A;Accession: JT0102
A;Molecule type: mRNA
A;Residues: 1-213,'I', 215-431 <NAG2>
A;Cross-references: GB:K03226; NID:g340155; PIDN:AAC97138.1; PID:g340158
R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RN
A;Reference number: A37561; MUID:84272706; PMID:6580620
A;Accession: A37561
A;Molecule type: mRNA
A;Residues: 66-431 <VER>
A;Cross-references: GB:D00244; NID:g220138
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen,
DNA 4, 139-146, 1985
A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prep
A;Reference number: I38102; MUID:85203359; PMID:3888571
A;Accession: I38102
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-150,'W',152-213,'I',215-385,'C',387-429,'V',431 <JAC>
A;Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298
R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki,
Biochim. Biophys. Acta 129, 83-89, 1996
A;Title: Characterization of single chain urokinase-type plasminogen activator with a no
A;Reference number: S65783; MUID:96186279; PMID:8652631
A;Accession: S65783
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 21-140,'L',142-213,'I',215-431 <YOS>
A;Cross-references: EMBL:D11143; NID:g1311467; PIDN:BAA01919.1; PID:g1199928
R;Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A;Title: The primary structure of high molecular mass urokinase from human urine.
A;Reference number: A37562; MUID:83055084; PMID:6754569
A;Accession: A37562
A;Molecule type: protein
A;Residues: 21-177 <GUN>
R;Schaller, J.; Nick, H.; Rickli, E.E.; Gillessen, D.; Lergier, W.; Studer, R.O.
Eur. J. Biochem. 125, 251-257, 1982
A;Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel
A;Reference number: A37563; MUID:83003608; PMID:6749491
A;Accession: A37563
A;Molecule type: protein
A;Residues: 156-176;179-193,'T',195,'T',197-224 <SCH>
R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A;Title: The complete amino acid sequence of low molecular mass urokinase from human uri
A;Reference number: A37564; MUID:83055099; PMID:6754572
A;Accession: A37564
A;Molecule type: protein
A;Residues: 158-410 <STE>
R;Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
Biochem. Biophys. Res. Commun. 171, 401-406, 1990

```
A;Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinar
A;Reference number: A35689; MUID:90365737; PMID:2393398
A;Accession: A35689
A;Molecule type: protein
A;Residues: 21-30,'X',32,'X',34-38,'X',40-43 <KEN>
A;Note: identification of a fucose and attempt to determine its attachment site
R;Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm
Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
A;Reference number: A36697; MUID:91097529; PMID:2125213
A;Accession: A36697
A;Molecule type: protein
A;Residues: 21-34 <RAB>
R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
submitted to the Brookhaven Protein Data Bank, July 1993
A;Reference number: A51255; PDB:1KDU
R;Li, X.; Smith, R.A.G.; Dobson, C.M.
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
Biochemistry 31, 9562-9571, 1992
A;Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
A;Reference number: A44375; MUID:93003110; PMID:132718
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR.
R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak,
submitted to the Brookhaven Protein Data Bank, January 1994
A;Reference number: A66822; PDB:1URK
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
submitted to the Brookhaven Protein Data Bank, July 1995
A;Reference number: A66058; PDB:1LMW
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175;179-426
C;Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
C;Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
C;Genetics:
A;Gene: GDB:PLAU
A;Cross-references: GDB:119497; OMIM:191840
A;Map position: 10q24-10q24
A;Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
A;Function:
A;Description: proteolytically activates plasminogen
A;Pathway: fibrinolysis
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteina
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predic
F;21-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <M
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <ME
F;179-419/Domain: trypsin homology <TRY>
F;31-39,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-40
F;38/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F;224,275,376/Active site: His, Asp, Ser #status experimental
F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match      37.7%;  Score 777; DB 1; Length 431;
Best Local Similarity  43.3%;  Pred. No. 1.4e-59;
Matches 157;  Conservative 48;  Mismatches 134;  Indels 24; Gaps 6;

Qy  29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
        ||| ||| ||| |||   || |||||   :: : | ||||||||||||||||||||
Db  70 CYEGNGHFYRGKASTDTMGRSCLPWNSATVLQQTYHAHRSDALQLGLGKHNYCRNPDNRR 129

Qy  89 KPWCHVLKNRRLTWEYCDVPSCS----------TCGLRQYSQPQFRIKGGLFADIA 134
        ||||: |  |  |:||  |  |||:            ::|:
Db 130 RPWCYVQVGLKPLVQECWVHDCADGKKPSSPPELKFQCG-QKTLRPRFKIIGGEFTTIE 188

Qy 135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
        : |||||  ||  ||||   |::| |||||||||:: :| ::|||||:|| ::|||||:
Db 189 NQPWFAAIYRRH-RGGSVTYVCGGSLMSPCWVISATHCFIDYPKKEDYIVYLGRSRLNSN 247

Qy 195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQL 252

Db 248 TQGEMKFEVENLIlHKDYSADTLAHHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDP 307
        :|||||  |||||| ||:  ||:||||||| : ||   |||:  || || |||
Qy 253 PDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRS 312

Db 308 QFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECQQPHYGSEVTTKMLCAAD--- 364
        ||| ||| | |||||  ||| | | ||| ||  | | ||||||||||||||
Qy 313 GGFQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYKVTNYLDWIRD 372

Db 373 NMR 375

Db 422 HTK 424

RESULT 11
UKBAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C;Accession: S14687; S08651
R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminoge
A;Reference number: S14687; MUID:90287734; PMID:2113276
A;Accession: S14687
A;Molecule type: mRNA
A;Residues: 1-433 <AUY>
A;Cross-references: EMBL:X51935; NID:g38130; PIDN:CAA36200.1; PID:g38131
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-176/Product: plasminogen activator chain A #status predicted <ACH>
F;30-61/Domain: EGF homology <EGF>
F;69-150/Domain: kringle homology <KRG>
F;178-433/Product: plasminogen activator chain B #status predicted <BCH>
F;178-421/Domain: trypsin homology <TRY>
F;167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;223,274,378/Active site: His, Asp, Ser #status predicted
F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      37.6%;  Score 774.5;  DB 1; Length 433;
Best Local Similarity  43.5%;  Pred. No. 2.3e-59;
Matches 160;  Conservative 45;  Mismatches 132;  Indels 31; Gaps 8;

Qy  29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
Db  69 CYEGNGHFYRGKASTDTMGRSCLAWNSATVLQQTYHAHRSDALQLGLGKHNYCRNPDNRR 128

Qy  89 KPWCHVLKNRRLTWEYCDVPSCS----------TCGLRQYSQPQFRIKGGLFADIA 134
Db 129 RPWCYVQVGLKQRVQECMVHNCADGKKPSSPPELQFQCGQRTL-RPRFKIVGGEFTTIB 187

Qy 135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
Db 188 NQPWFAAIYRRH-RGGSVTYVCGGSLISPCWVVSATHCFINYPKKEDYIVYLGRSRLNSN 246

Qy 195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLP-----P 247
Db 247 TQGEMKFEVENLIHEDYSADTLAHHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDP 306

Qy 248 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 307
Db 307 ND--PPFGTSCEITGFGKENSTDYLYPEQLKMTVVKLVSHQKCQQPHYYGSEVTTKMLCA 364

Qy 308 GDTRSGGFQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYVL 367
Db 365 AD------PQWET-DSCQGDSGGPLVCSIQGHMTLTGIVSWGRGCALKDKPGVTRVSRFL 418

Qy 368 DWIRDNMR 375
Db 419 PWIHSHTR 426
```

RESULT 12
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N;Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
C;Accession: S24604; I60186; I53472; S18932
R;Rabbani, S.A.
A;Reference number: S24604
A;Accession: S24604
A;Molecule type: mRNA
A;Residues: 1-15,'H',17-23,'G',25-331,'N',333-432 <RAB>
A;Cross-references: EMBL:X65651; NID:g574456; PIDN:CAA46601.1; PID:g57457
A;Experimental source: tissue kidney
R;Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
A;Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
Cancer Res. 52, 2489-2496, 1992
A;Reference number: I60186; MUID:92233409; PMID:1568219
A;Accession: I60186
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-432 <RES>
A;Cross-references: EMBL:X63434; NID:g57465; PIDN:CAA45028.1; PID:g57466
A;Experimental source: strain Fischer 344; tissue mammary
R;Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A;Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A;Reference number: I53472; MUID:92339549; PMID:1321734
A;Accession: I53472
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 31-62 <RE2>
A;Cross-references: EMBL:X66907; NID:g396200; PIDN:CAA47356.1; PID:g938279
C;Genetics:
A;Gene: uPA
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;179-420/Domain: trypsin homology <TRY>
F;168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted
F;225,276,377/Active site: His, Asp, Ser #status predicted

Query Match   37.5%; Score 772.5; DB 1; Length 432;
Best Local Similarity   40.6%; Pred. No. 3.5e-59;
Matches 147; Conservative 64; Mismatches 128; Indels 23; Gaps 5;

```
Qy   29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
         || ||| : ||   |    | ||| ||  :| | ||| ||::      |||||||||||
Db   70 CYHGNGQSYRGKANTDTKGRPCLAWNSPAVLQQTYNAHRSDALSLGLGKHNYCRNPDNQR 129

Qy   89 KPWCHVLKNRRLTWEYCDVPSCS---------------TCGLRQYSQPQFRIKGGLFADIA 134
         :||| :  | :| | || :|| :               ||||:| :::|  |||  :
Db  130 RPWCYVQIGLKQFVQECMVQDCSLSKKPSSTVDQQGFQCG--QKALRPRFKIVGGEFTVVE 188

Qy  135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
         : ||:AAI: |: | ::  ||  :|| :||| :||:|||:   :  :  | |
Db  189 NQPWFAAIYLKNKGGSPPSFKCGGSLISPCWVASATHCFVNQPKKEEVVYIGQSKRNSY 248

Qy  195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQL 252
          :  :KFEVE :|:||:| : : : ||||||::: : : ||| : : ||:|||| : :
Db  249 NPGEMKFEVEQLILHEDFSDETLAFHNDIALLKIRTSTGQCAQPSRTIQTICLPPRFGDA 308

Qy  253 PDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRS 312
            | |:: | || :|  ||:|:| ||  : :| | |:|  :: |  : |||||  :
Db  309 PFGSDCEITGFGQESATDYFYPKDLKMSVVKIISHEQCCKQPHYYGSEINYKMLCAAD---365

Qy  313 GGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRD 372
         ||| :   :|| |||||||:| | || || || |||  || :| ||||||| :: :|| :
Db  366 --PEWKT-DSCSGDSGGPLICNIDGRPTLSGIVSWGSGCAEKNKPGVYTRVSYFLNWIQS 422

Qy  373 NM 374
         ::
Db  423 HI 424
```

RESULT 13
UKPG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N;Alternate names: uPA
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C;Accession: A00932
R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A;Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A;Reference number: A00932; MUID:85087954; PMID:6096832
A;Accession: A00932
A;Molecule type: DNA
A;Residues: 1-240,'H',242-442 <NAG1>
A;Experimental source: kidney cell line LLC-PK1
R;Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A;Reference number: A37566
A;Contents: annotation; correction to residue 241
C;Genetics:
A;Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;190-430/Domain: trypsin homology <TRY>
F;152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;179-310,220-228,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
F;235,286,387/Active site: His, Asp, Ser #status predicted

Query Match   37.3%; Score 769.5; DB 1; Length 442;
Best Local Similarity   41.2%; Pred. No. 6.5e-59;
Matches 153; Conservative 50; Mismatches 135; Indels 33; Gaps 6;

```
Qy   29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
         || |||:|||  | ||   | ||||| :|::  ||||  ||:   |||||||||||| |:
Db   72 CFEGNGHSYRGKANTNTGGRPCLPWNSATVLLNTYHAHRPDALQLGLGKHNYCRNPDNQR 131

Qy   89 KPWCHVLKNRRLTWEYCDVPSCS---------------TCGLRQYSQPQPRI 125
         :||| :  | :  | ||:| ||                 ||:| :  | :|:
Db  132 RPWCYVQVGLKQLVQECMVPNCSGGSHRPAYDGKNPFSTPBKVEFQCG-QKALRPRFKI 190

Qy  126 KGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVI 185
         :||      : ||:AAI: ::  :   : | |: || |||:::|:||||:   : : :|
Db  191 VGGKSTTIENQPWFAAIYRRH-RGGSVTYVCGGSLISPCWVVSATHCFINYQQKEDYIVY 249

Qy  186 LGRTYRVVPGEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTV 243
         |||   : : : :KFEVEK :::||:: : :   ||||||::: | : ||| | : :
Db  250 LGRQTLHSSTHGMKFEVEKLILHEDYSADSLAHHNDIALLKIRTDKGQCAQPSRSIQTI 309

Qy  244 CLPPADLQLPDWTECELSGYGHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDN 303
         ||||   :    | |::  ||::   : :|: || : |:| | |:| :: :    :: |
Db  310 CLPPVNGDAHFGASCEIVGFGKEDPSDYLYPEQLKMTVVKLVSHRECQQPHYYGSEVTTK 369

Qy  304 MLCAGDTRSGGPGQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKV 363
         |||| |      ::      ||||||||||| : || || || |||  |: ||||||| :|
Db  370 MLCAAD------PQWKT-DSCQGDSGGPLVCSTQGRLTLTGIVSWGRECAMKDKPGVYTRV 423

Qy  364 TNYLDWIRDNM 374
          :: ||| :
Db  424 SRFLTWIHTHV 434
```

RESULT 14
A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N;Alternate names: uPA
C;Species: Gallus gallus (chicken)
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C;Accession: A35005
R;Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A;Title: The chicken urokinase-type plasminogen activator gene.
A;Reference number: A35005; MUID:90110185; PMID:2295632
A;Accession: A35005
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-434 <LES>
A;Cross-references: GB:J05187; NID:g212858; PIDN:AAA49131.1; PID:g212859
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;40-79/Domain: EGF homology <EGF>
F;79-158/Domain: kringle homology <KRG>
F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;173-421/Domain: trypsin homology <TRY>
F;162-296,210-218,285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F;217,272,373/Active site: His, Asp, Ser #status predicted

Query Match   37.0%; Score 762; DB 1; Length 434;
Best Local Similarity   45.4%; Pred. No. 2.9e-58;
Matches 163; Conservative   56; Mismatches 116; Indels 24; Gaps 12;

```
Qy   26  NSDCYFGNGSAYRGTHSLTESGASCLPWN--SMILIGKVYTAQNPSAQALGLGKHNYCRN  83
         ||| ||  ||   ||       |||  ||   : ||   |||:|:||||||||||||||||
Db   76  NSICYSGNGEDYRG----MAEDPGCLYWDHPSVIRWGD-YHADLKNALQLGLGKHNYCRN 130

Qy   84  PDGDAKPWCHVLKNRRLTWEYCD-VPSCS-TCGLRQYSQPQFRIKGGLFADIASHPWQAA 141
         | || :|||  :||||   : |   | || |  | | ||  | :|||| ::| || ||||
Db  131  PNGRSRPWCYTKRRYSIQETPCSTIEKCERTCGQRSFSK-YFKIVGGSQAEVETQPWIAG 189

Qy  142  IFAKHRRSPGERFLCGGILISSCWILSAAHCFQE--RFPPHH--LTVILGRTYRVVPGEE 197
         ||   :   |  ||||| || ||| |||||||:     | ||    :| |::: :  : :
Db  190  IFQNIMGT--DQFLCGGSLIDPCWVLTAAHCFYNPTKKQPNKSVYKVFLGKSILNTNDEH 247

Qy  198  EQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQBSSVRRTVCLPPADLQLPDW 255
         ||:|   :  |:: |  ||  ||||||| : ::  : ||  || |||||||  :| | | :
Db  248  EQVFMVDEIISHPDFTDHTGGNDNDIALIRIRTASGQCAVESNYVRTVCLPEKNLNLYDN 307

Qy  256  TECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGP 315
         | |: :|||:    | ::  :|: | | : | :::| :  : ::::|||:||||     |
Db  308  TWCEIAGYGKQNSYDIYYAQRLMSATVNLISQDDCKNKYYDSTRVTDNMVCAGD-----P 362

Qy  316  QANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNM 374
         ||| | || |||||| ||  :|||||| |:|||:| | : ||||||:||  ||:| :||
Db  363  LWET-DACKGDSGGPMVCEHNGRMTLYGIVSWGNGDGCAKKNKPGVYTRVTRYLNWIDSNM 420
```

RESULT 15
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N;Alternate names: uPA
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: JN0560
R;Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
A;Reference number: JN0560; MUID:93216119; PMID:8385052
A;Accession: JN0560
A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:L03546; NID:g163800; PIDN:AAA51419.1; PID:g163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA1>
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;181-421/Domain: trypsin homology <TRY>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match   36.1%; Score 745; DB 1; Length 433;
Best Local Similarity   41.3%; Pred. No. 8.6e-57;
Matches 150; Conservative   54; Mismatches 135; Indels 24; Gaps   6;

```
Qy   29  CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA  88
         || ||  :||  |     | |||  :  :||   |||:||||||||||||||||||||:
Db   72  CYQGNGHSYRGKANRDLSGRPCLAWDSPTVLLKMYHAHRSDAIQLGLGKHNYCRNPDNQR 131

Qy   89  KPWCHVLKNRRLTWEYCDVPSCST------------CGLRQYSQPQFRIKGGLFADIA   134
         :|||  : |:  : : |: | |             |||:|  :|: | :|||||| |
Db  132  RPWCYVQIGLKQPVQFCMVQDCSVGKSPSSPREKEEFQCG-QKALRPRFKIVGGQVTNAE 190

Qy  135  SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
          : ||  || : : || |  | |||||  || |||||||:   : ||  : : | : :
Db  191  NQPWFAAIYRRH-RGGSITYLCGGSLISPCWVSATHCFIDHPKKENYIVYLGQSRLNSD  249

Qy  195  GEBEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQBSSVVRTVCLPPADLQL 252
            ::||||||:::::| ::      :|||||| :::: : ||  ||: ||||| |::  |
Db  250  TRGEMQFEVEKLILHEDYSAESLAHHNDIALLKIRTSRGQCAQPSRSIQTICLPPEHEDA 309

Qy  253  PDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRS 312
          ::| | :|:||:   | ::|  : ::| | : : : ||   : ||::::||:||  :
Db  310  HSRTRCEITGFGKENPSDYRYSDELKMTFVSLLVSHEVCQQPHYYGAEVTDKMLCAAD--- 366

Qy  313  GGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRD 372
            ||| | || |||||||||| : |:||| ||:|||:  | :  ||||||:|:::| ::
Db  367  --PQWET-DSCQGDSGGPLVCTIQGRLTLTGIVSWGRDCAMKYKPGVYTRVSKFLPWINT 423

Qy  373  NMR 375
          :|
Db  424  HTR 426
```

Search completed: August 8, 2003, 17:15:24
Job time : 20 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August  8, 2003, 17:10:32 ; Search time 15 Seconds
                                        (without alignments)
                              1181.938 Million cell updates/sec

Title:          US-09-987-455-8
Perfect score:  2061
Sequence:       1 MKKTAIAIAVALAGFATVAQ..........GVYTKVTNYLDWIRDNMRPG 377

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:     127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SwissProt_41:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                         SUMMARIES

Result              Query
No.    Score   Match  Length  DB  ID           Description
------ -----   -----  ------  --  ----------   -----------
1      1961    95.1    562    1   TPA_HUMAN    P00750 homo sapien
2      1660    80.5    566    1   TPA_BOVIN    Q28198 bos taurus
3      1638.5  79.5    559    1   TPA_RAT      P19637 rattus norv
4      1601    77.7    559    1   TPA_MOUSE    P11214 mus musculu
5      1451.5  70.4    394    1   URTG_DESRO   P49150 desmodus ro
6      1428    69.3    477    1   URT2_DESRO   P15638 desmodus ro
7      1427    69.2    431    1   URTB_DESRO   P98121 desmodus ro
8      1389.5  67.4    477    1   URT1_DESRO   P98119 desmodus ro
9      779.5   37.8    433    1   UROK_MOUSE   P06869 mus musculu
10     777     37.7    431    1   UROK_HUMAN   P00749 homo sapien
11     774.5   37.6    433    1   UROK_PAPCY   P16227 papio cynoc
12     772.5   37.5    432    1   UROK_RAT     P29598 rattus norv
13     769.5   37.3    442    1   UROK_PIG     P04185 sus scrofa
14     762     37.0    434    1   UROK_CHICK   P15120 gallus gall
15     745     36.1    433    1   UROK_BOVIN   Q05589 bos taurus
16     742.5   36.0    653    1   HGFA_MOUSE   Q9r098 mus musculu
17     738     35.8    655    1   HGFA_HUMAN   Q04756 homo sapien
18     671.5   32.6    615    1   FA12_HUMAN   P00748 homo sapien
19     670     32.5    603    1   FA12_CAVPO   Q04962 cavia porce
20     620.5   30.1    593    1   FA12_BOVIN   P98140 bos taurus
21     607     29.5    812    1   PLMN_MOUSE   P20918 mus musculu
22     573.5   27.8    333    1   PLMN_CANFA   P80009 canis famil
23     571.5   27.7    343    1   PLMN_SHEEP   P81286 ovis aries
24     563.5   27.3    810    1   PLMN_MACMU   P12545 macaca mula
25     561.5   27.2    810    1   PLMN_HUMAN   P00747 homo sapien
26     552.5   26.8    790    1   PLMN_PIG     P06867 sus scrofa
27     550.5   26.7    812    1   PLMN_BOVIN   P06868 bos taurus
28     547     26.5    810    1   PLMN_ERIEU   Q29485 erinaceus e
29     541.5   26.3    4548   1   APOA_HUMAN   P08519 homo sapien
30     538.5   26.1    338    1   PLMN_HORSE   P80010 equus cabal
31     519     25.2    1420   1   APOA_MACMU   P14417 macaca mula
32     469.5   22.8    875    1   NETR_HUMAN   P56730 homo sapien
33     465     22.6    761    1   NETR_MOUSE   O08762 mus musculu
```

```
34     449     21.8    855    1   ST14_HUMAN   Q9y5y6 homo sapien
35     438     21.3    855    1   ST14_MOUSE   P56677 mus musculu
36     408     19.8    455    1   TMS5_MOUSE   Q9er04 mus musculu
37     404     19.6    490    1   TMS2_MOUSE   Q9jiq8 mus musculu
38     403     19.6    457    1   TMSS_HUMAN   Q9h3s3 homo sapien
39     402     19.5    418    1   HATT_HUMAN   O60235 homo sapien
40     399     19.4    603    1   CFAI_MOUSE   Q61129 mus musculu
41     398.5   19.3    275    1   TRYT_PIG     Q9n2d1 sus scrofa
42     398.5   19.3    625    1   THRB_BOVIN   P00735 bos taurus
43     396.5   19.2    261    1   KLK1_MOUSE   P15947 mus musculu
44     396.5   19.2    618    1   THRB_MOUSE   P19221 mus musculu
45     394     19.1    604    1   CFAI_RAT     Q9wuw3 rattus norv
```

                              ALIGNMENTS

```
RESULT 1
TPA_HUMAN
ID   TPA_HUMAN     STANDARD;      PRT;    562 AA.
AC   P00750; Q15103;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).
GN   PLAT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Melanoma;
RX   MEDLINE=83115262; PubMed=6637343;
RA   Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA   Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA   Goeddel D.V., Collen D.;
RT   "Cloning and expression of human tissue-type plasminogen activator
RT   cDNA in E. coli.";
RL   Nature 301:214-221(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal lung;
RX   MEDLINE=88262579; PubMed=3313640;
RA   Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT   "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT   from human fetal lung cells.";
RL   Nucleic Acids Res. 16:5695-5695(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88054470; PubMed=2824147;
RA   Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA   Hsiung N.;
RT   "Expression of human uterine tissue-type plasminogen activator in
RT   mouse cells using BPV vectors.";
RL   DNA 6:461-472(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86196143; PubMed=3009482;
RA   Friezner Degen S.J., Rajput B., Reich E.;
RT   "The human tissue plasminogen activator gene.";
RL   J. Biol. Chem. 261:6972-6985(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84298137; PubMed=6089198;
RA   Ny T., Elgh F., Lund B.;
RT   "The structure of the human tissue-type plasminogen activator gene:
RT   correlation of intron and exon structures to functional and
RT   structural domains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN   [6]
RP   SEQUENCE FROM N.A.
```

RX   MEDLINE=86284200; PubMed=3090401;
RA   Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA   Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT   "Cloning of cDNA coding for human tissue-type plasminogen activator
RT   and its expression in Escherichia coli.";
RL   Mol. Biol. Med. 3:279-292(1986).
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Umbilical vein;
RX   MEDLINE=90192129; PubMed=2107528;
RA   Siebert P.D., Fong K.;
RT   "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT   human endothelial cells.";
RL   Nucleic Acids Res. 18:1086-1086(1990).
RN   [8]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [9]
RP   SEQUENCE OF 212-361 FROM N.A.
RX   MEDLINE=83169656; PubMed=6572897;
RA   Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA   Josephson S.;
RT   "Isolation of cDNA sequences coding for a part of human tissue
RT   plasminogen activator.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN   [10]
RP   SEQUENCE OF 1-36 FROM N.A.
RX   MEDLINE=85289338; PubMed=3161893;
RA   Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
RA   Schleuning W.-D.;
RT   "Isolation and characterization of the human tissue-type plasminogen
RT   activator structural gene including its 5' flanking region.";
RL   J. Biol. Chem. 260:11223-11230(1985).
RN   [11]
RP   SEQUENCE OF 31-562 FROM N.A.
RX   MEDLINE=91291340; PubMed=1368681;
RA   Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT   "Purification and characterization of tissue plasminogen activator
RT   secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
RL   Agric. Biol. Chem. 55:1225-1232(1991).
RN   [12]
RP   SEQUENCE OF 36-562.
RC   TISSUE=Melanoma;
RX   MEDLINE=85000468; PubMed=6433976;
RA   Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT   "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT   derived amino acid sequence, identify the active site serine residue,
RT   establish glycosylation sites, and localize variant differences.";
RL   Biochemistry 23:3701-3707(1984).
RN   [13]
RP   SEQUENCE OF 33-52 AND 311-330.
RC   TISSUE=Melanoma;

RX   MEDLINE=83209620; PubMed=6682760;
RA   Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT   "Purification and characterization of a melanoma cell plasminogen
RT   activator.";
RL   Eur. J. Biochem. 132:681-686(1983).
RN   [14]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=90092112; PubMed=2513186;
RA   Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT   "Carbohydrate structure of recombinant human uterine tissue
RT   plasminogen activator expressed in mouse epithelial cells.";
RL   Eur. J. Biochem. 186:273-286(1989).
RN   [15]
RP   CARBOHYDRATE-LINKAGE SITE THR-96.
RX   MEDLINE=91159408; PubMed=1900431;
RA   Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT   "Tissue plasminogen activator has an O-linked fucose attached to
RT   threonine-61 in the epidermal growth factor domain.";
RL   Biochemistry 30:2311-2314(1991).
RN   [16]
RP   DISULFIDE BONDS IN KRINGLE 2.
RX   MEDLINE=91244765; PubMed=1645336;
RA   Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT   "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT   plasminogen activator produced in Escherichia coli.";
RL   J. Biol. Chem. 266:10070-10072(1991).
RN   [17]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX   MEDLINE=96200985; PubMed=8613982;
RA   Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA   Bode W.;
RT   "The 2.3 A crystal structure of the catalytic domain of recombinant
RT   two-chain human tissue-type plasminogen activator.";
RL   J. Mol. Biol. 258:117-135(1996).
RN   [18]
RP   X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX   MEDLINE=97449126; PubMed=905622;
RA   Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA   Bode W.;
RT   "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT   crystal structure of single-chain human tPA.";
RL   EMBO J. 16:4797-4805(1997).
RN   [19]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX   MEDLINE=92118803; PubMed=1310033;
RA   de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
RA   Westbrook M.L., Kossiakof A.A.;
RT   "Crystal structure of the kringle 2 domain of tissue plasminogen
RT   activator at 2.4 A resolution.";
RL   Biochemistry 31:270-279(1992).
RN   [20]
RP   STRUCTURE BY NMR OF KRINGLE 2.
RX   MEDLINE=90122799; PubMed=2558718;
RA   Byeon I.-J.L., Kelley R.F., Llinas M.;
RT   "1H NMR structural characterization of a recombinant kringle 2 domain
RT   from human tissue-type plasminogen activator.";
RL   Biochemistry 28:9350-9360(1989).
RN   [21]
RP   STRUCTURE BY NMR OF KRINGLE 2.
RX   MEDLINE=91200042; PubMed=1901789;
RA   Byeon I.-J.L., Kelley R.F., Llinas M.;
RT   "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
RT   assignments and secondary structure.";
RL   Eur. J. Biochem. 197:155-165(1991).
RN   [22]
RP   STRUCTURE BY NMR OF KRINGLE 2.
RX   MEDLINE=92106329; PubMed=1762144;
RA   Byeon I.-J.L., Llinas M.;
RT   "Solution structure of the tissue-type plasminogen activator kringle
RT   2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
RT   drug.";
RL   J. Mol. Biol. 222:1035-1051(1991).
RN   [23]

```
Query Match   95.1%;  Score 1961; DB 1; Length 562;
Best Local Similarity  99.7%;  Pred. No. 1.1e-162;
Matches  355;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY  80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  267 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 326

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 386

Qy  201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  387 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 446

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  447 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 506

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGCQKDVPGVYTKVTNYLDWIRDNMRP 562


RESULT 2
TPA_BOVIN
ID   TPA_BOVIN      STANDARD;      PRT;    566 AA.
AC   Q28198;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   PLAT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Ravn P., Berglund L., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen activators uPA
RT   and tPA.";
RL   Int. Dairy J. 5:605-617(1995).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC       BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC       PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC       ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC       CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; X85800; CAA59795.1; -.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctn1.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS01185; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT   SIGNAL        1     21
FT   PROPEP       22     33       BY SIMILARITY.
FT   CHAIN        34    566       TISSUE-TYPE PLASMINOGEN ACTIVATOR A.
FT   CHAIN        34    314       TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT                                CHAIN.
FT   CHAIN       315    566       TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT                                CHAIN.
FT   DOMAIN       40     82       FIBRONECTIN TYPE-I.
FT   DOMAIN       83    121       EGF-LIKE.
FT   DOMAIN      128    209       KRINGLE 1.
FT   DOMAIN      219    300       KRINGLE 2.
FT   DOMAIN      315    566       SERINE PROTEASE.
FT   ACT_SITE    361    361       CHARGE RELAY SYSTEM.
FT   ACT_SITE    410    410       CHARGE RELAY SYSTEM.
FT   ACT_SITE    517    517       CHARGE RELAY SYSTEM.
FT   DISULFID     42     72       BY SIMILARITY.
FT   DISULFID     70     79       BY SIMILARITY.
FT   DISULFID     87     98       BY SIMILARITY.
FT   DISULFID     92    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    128    209       BY SIMILARITY.
FT   DISULFID    149    191       BY SIMILARITY.
FT   DISULFID    180    204       BY SIMILARITY.
FT   DISULFID    219    300       BY SIMILARITY.
FT   DISULFID    240    282       BY SIMILARITY.
FT   DISULFID    271    295       BY SIMILARITY.
FT   DISULFID    303    434       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    346    362       BY SIMILARITY.
FT   DISULFID    354    423       BY SIMILARITY.
FT   DISULFID    448    523       BY SIMILARITY.
FT   DISULFID    480    496       BY SIMILARITY.
FT   DISULFID    513    541       BY SIMILARITY.
FT   CARBOHYD    153    153       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    487    487       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   566 AA;  63701 MW;  2EB6BBB4E32276C3 CRC64;
```

```
Query Match    80.5%;  Similarity  82.8%;  Score 1660;  DB 1;  Length 566;
Best Local Similarity    82.8%;  Pred. No. 1.5e-136;
Matches 293;  Conservative   24;  Mismatches   37;  Indels    0;  Gaps    0;

Qy     23  SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQALGLGKHNYCR   82
               :| |||||||||||:|||||||||||||||||: :|||:  :  :|:||| |||||||||
Db    213  AEBDGCYTGNGLAYRGRSRHTKSGASCLPWNSVFLTSKIYTAWKSNAPALGLGKHNHCR  272

Qy     83  NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI  142
               |||||:||| ||||||||||||||||||||:|||||||:|||||||||||||||||||||
Db    273  NPDGDAQPWCHVWKDRQLTWEYCDVPQCVTCGLRQYRPQFRIKGGLFADITSHPWQAAI   332

Qy    143  FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFE  202
               | :||||||||||||||||:||||||||||||||||:|||||||||||:||||||| |||
Db    333  FVKNRRSPGERFLCGGILISSCWVLSAAHCFQERYPPHHLKVFLGRTYRLVPGEEBQTFE  392

Qy    203  VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG  262
               |||||||||||||||||||||||||:||:|| ||||:|||||||||:||||||||||||||
Db    393  VEKYIIHKEFDDDTYDNDIALLHLKSDSLTCARESASVRTICLPDASLQLPDWTECELSG  452

Qy    263  YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  322
               ||||| :||||||||||||||||||||||||||:||||:||||||||||||||      |||||
Db    453  YGKHESSSPFFSERLKEAHVRLYPSSRCTSQHLFNRTVTNNMLCAGDHTRSGGDHTNLHDA  512

Qy    323  CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP  376
               |||||||||||||:|| |:||||||||||||:|||||||||||||||||||||||||
Db    513  CQGDSGGPLVCMKDNHMTLVGIISWGLGCGRKDVPGVYTKVTNYLDWIRDNTRP  566
```

```
RESULT 3
TPA_RAT
ID   TPA_RAT      STANDARD;      PRT;    559 AA.
AC   P19637;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   PLAT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89170114; PubMed=3148445;
RA   Ny T., Leonardsson G., Hueh A.J.W.;
RT   "Cloning and characterization of a cDNA for rat tissue-type
RT   plasminogen activator.";
RL   DNA 7:671-677(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90130448; PubMed=2105315;
RA   Feng P., Ohlsson M., Ny T.;
RT   "The structure of the TATA-less rat tissue-type plasminogen activator
RT   gene. Species-specific sequence divergences in the promoter predict
RT   differences in regulation of gene expression.";
RL   J. Biol. Chem. 265:2022-2027(1990).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC       BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC       PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC       ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC       CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M23697; AAA41812.1; -.
DR   EMBL; M31197; AAA42261.1; JOINED.
DR   EMBL; M31185; AAA42261.1; JOINED.
DR   EMBL; M31186; AAA42261.1; JOINED.
DR   EMBL; M31187; AAA42261.1; JOINED.
DR   EMBL; M31188; AAA42261.1; JOINED.
DR   EMBL; M31189; AAA42261.1; JOINED.
DR   EMBL; M31190; AAA42261.1; JOINED.
DR   EMBL; M31191; AAA42261.1; JOINED.
DR   EMBL; M31192; AAA42261.1; JOINED.
DR   EMBL; M31193; AAA42261.1; JOINED.
DR   EMBL; M31194; AAA42261.1; JOINED.
DR   EMBL; M31195; AAA42261.1; JOINED.
DR   EMBL; M31196; AAA42261.1; JOINED.
DR   EMBL; A19618; CAA01482.1; -.
DR   PIR; A35029; A35029.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctn1.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Plasma; Kringle; EGF-like domain; Repeat; Signal;
FT   SIGNAL         1     17
FT   PROPEP        18     29
FT   CHAIN         30    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN         30    308       TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT                                 CHAIN.
FT   CHAIN        309    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT                                 CHAIN.
FT   DOMAIN        36     78       FIBRONECTIN TYPE-I.
FT   DOMAIN        79    117       EGF-LIKE.
FT   DOMAIN       124    205       KRINGLE 1.
FT   DOMAIN       213    294       KRINGLE 2.
FT   DOMAIN       309    559       SERINE PROTEASE.
FT   ACT_SITE     355    355       CHARGE RELAY SYSTEM.
FT   ACT_SITE     404    404       CHARGE RELAY SYSTEM.
FT   ACT_SITE     510    510       CHARGE RELAY SYSTEM.
```

```
FT   DISULFID      38     68       BY SIMILARITY.
FT   DISULFID      66     75       BY SIMILARITY.
FT   DISULFID      83     94       BY SIMILARITY.
FT   DISULFID      88    105       BY SIMILARITY.
FT   DISULFID     107    116       BY SIMILARITY.
FT   DISULFID     124    205       BY SIMILARITY.
FT   DISULFID     145    187       BY SIMILARITY.
FT   DISULFID     176    200       BY SIMILARITY.
FT   DISULFID     213    294       BY SIMILARITY.
FT   DISULFID     234    276       BY SIMILARITY.
FT   DISULFID     265    289       BY SIMILARITY.
FT   DISULFID     297    428       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     340    356       BY SIMILARITY.
FT   DISULFID     348    417       BY SIMILARITY.
FT   DISULFID     442    516       BY SIMILARITY.
FT   DISULFID     474    490       BY SIMILARITY.
FT   DISULFID     506    534       BY SIMILARITY.
FT   CARBOHYD     149    149       N-LINKED (GLCNAC...  .) (POTENTIAL).
FT   CARBOHYD     481    481       N-LINKED (GLCNAC...  .) (POTENTIAL).
FT   CONFLICT     380    380       E -> K (IN REF. 1).
SQ   SEQUENCE   559 AA;  62903 MW;  7DBD3809C1D1C921 CRC64;

Query Match             79.5%;  Score 1638.5;  DB 1;  Length 559;
Best Local Similarity   82.5%;  Pred. No. 1.1e-134;
Matches  287;  Conservative   29;  Mismatches   31;  Indels    1;  Gaps    1;

Qy         28 DCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGD 87
              |||||||| ||||||| :: |||||||| :: |||||||:|||||||:|||||||||||||
Db        212 DCYVGKGVTYRGTHSFTTSKASCLPWNSMILIGKTYTAWRANSQALGLGRHNYCRNPDGD 271

Qy         88 AKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHR 147
              |||||||::|:: |||||||||||||||  ||||||||| || || ||||||||  |:::
Db        272 AKPWCHVMKDRKLTWEYCDMSPSCSTCGLRQYKQPQFRIKGGLFTDITSHPWQAAIFVKNK 331

Qy        148 RSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFEVEKYI 207
              ||||||||||||||||||||||||||||| || |||||||||||||||||| ::||||||
Db        332 RSPGERFLCGGVLISSCWVLSAAHCFVERFPPHHLKVVLGRTYRVVPGEEEQTFEIEKYI 391

Qy        208 VHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHE 267
              |||||||||||||||||||||||:|||:||||| | ||||||||||||||||||||||
Db        392 VHKEFDDDTYDNDIALLQLRSDSSQCAQESSVGTACLLPDPDVQLPDWTECELSGYGKHE 451

Qy        268 ALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACOGDS 327
              ||:|| ::|||||||:|||||||||||:||::|: |:|||||||| ::||||||||||||
Db        452 ASSPFFSDRLKEAHVRLYPSSRCTSQHLFNKTITSNMLCAGDTRGGNQ-DVHDACQGDS 510

Qy        328 GGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRDNMR 375
              ||||:||| ||| ||||||||||||| |||| ||||||||||||:|||
Db        511 GGPLVCMIDKRMTLLGIISWGLGCGQKDVPGIYTKVTNYLNWIQDNMK 558

RESULT 4
TPA_MOUSE
ID   TPA_MOUSE      STANDARD;      PRT;   559 AA.
AC   P11214; Q91VP2;
DT   01-JUL-1989 (Rel. 11, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   PLAT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88087303; PubMed=2826484;
RA   Rickles R.J., Darrow A.L., Strickland S.;
RT   "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT   activator mRNA and its expression during F9 teratocarcinoma cell
RT   differentiation.";
RL   J. Biol. Chem. 263:1563-1569(1988).

RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC       TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC       CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC       ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC       MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC       BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC       PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC       ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC       CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb.ch).
DR   EMBL; J03520; AAA40470.1; -.
DR   EMBL; BC011256; AAH11256.1; -.
DR   PIR; A29941; A29941.
DR   HSSP; P00750; 1A5H.
DR   MEROPS; S01.232; -.
DR   MGD; MGI:97610; Plat.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctnl.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
```

```
Qy    320 HDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMR 375
          |||||||||||||| |||:.: |||| || ::||||||||||||||||||||||||  ||:
Db    503 HDACQGDSGGPLVCMINKQMTLTGIISWGLGCGQKDVPGVYTKVTNYLDWIHDNMK 558
```

```
RESULT 5
ID  URTG_DESRO     STANDARD;      PRT;      394 AA.
AC  P49150;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE  gamma).
OS  Desmodus rotundus (Vampire bat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC  Desmodontinae; Desmodus.
OX  NCBI_TaxID=9430;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Salivary gland;
RX  MEDLINE=92039036; PubMed=1937019;
RA  Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA  Alagon A., Donner P., Schleuning W.D.;
RT  "The plasminogen activator family from the salivary gland of the
RT  vampire bat Desmodus rotundus: cloning and expression.";
RL  Gene 105:229-237(1991).
RN  [2]
RP  CHARACTERIZATION.
RX  MEDLINE=93393059; PubMed=1309059;
RA  Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA  Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA  Donner P.;
RT  "Plasminogen activators from the saliva of Desmodus rotundus (common
RT  vampire bat): unique fibrin specificity.";
RL  Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC  -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC      EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC      AGENT.
CC  -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC      plasminogen to form plasmin.
CC  -!- SUBUNIT: Monomer.
CC  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC  -!- SIMILARITY: Contains 1 kringle domain.
CC  --------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
DR  EMBL; M63990; AAA31595.1; -.
DR  PIR; JS0600; JS0600.
DR  HSSP; P98119; 1A5I.
DR  MEROPS; S01.239; -.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR001254; Ser_protease_Try.
DR  Pfam; PF00051; kringle; 1.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 1.
DR  SMART; SM00130; KR; 1.
DR  SMART; SM00020; Tryp_SPc; 1.
DR  PROSITE; PS00021; KRINGLE_1; 1.
DR  PROSITE; PS50070; KRINGLE_2; 1.
DR  PROSITE; PS00240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
```

Second column:

```
DR  SMART; SM00130; KR; 2.
DR  SMART; SM00020; Tryp_SPc; 1.
DR  PROSITE; PS00022; EGF_1; 1.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS01253; FIBRONECTIN_1; 1.
DR  PROSITE; PS00021; KRINGLE_1; 2.
DR  PROSITE; PS50070; KRINGLE_2; 2.
DR  PROSITE; PS50240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
KW  Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW  Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT  SIGNAL        1     17
FT  PROPEP       18     29       PROBABLE.
FT  CHAIN        30    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT  CHAIN        30    308       TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT                               CHAIN.
FT  CHAIN       309    559       TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT                               CHAIN.
FT  DOMAIN       36     78       FIBRONECTIN TYPE-I.
FT  DOMAIN       79    117       EGF-LIKE.
FT  DOMAIN      124    205       KRINGLE 1.
FT  DOMAIN      213    294       KRINGLE 2.
FT  DOMAIN      309    559       SERINE PROTEASE.
FT  ACT_SITE    355    355       CHARGE RELAY SYSTEM.
FT  ACT_SITE    404    404       CHARGE RELAY SYSTEM.
FT  ACT_SITE    510    510       CHARGE RELAY SYSTEM.
FT  DISULFID     38     68       BY SIMILARITY.
FT  DISULFID     66     75       BY SIMILARITY.
FT  DISULFID     83     94       BY SIMILARITY.
FT  DISULFID     88    105       BY SIMILARITY.
FT  DISULFID    107    116       BY SIMILARITY.
FT  DISULFID    124    205       BY SIMILARITY.
FT  DISULFID    145    187       BY SIMILARITY.
FT  DISULFID    176    200       BY SIMILARITY.
FT  DISULFID    213    294       BY SIMILARITY.
FT  DISULFID    234    276       BY SIMILARITY.
FT  DISULFID    265    289       BY SIMILARITY.
FT  DISULFID    297    428       INTERCHAIN (BY SIMILARITY).
FT  DISULFID    340    356       BY SIMILARITY.
FT  DISULFID    348    417       BY SIMILARITY.
FT  DISULFID    442    516       BY SIMILARITY.
FT  DISULFID    474    490       BY SIMILARITY.
FT  DISULFID    506    534       BY SIMILARITY.
FT  CARBOHYD    149    149       N-LINKED (GLCNAC...)  (POTENTIAL).
FT  CARBOHYD    481    481       N-LINKED (GLCNAC...)  (POTENTIAL).
FT  CONFLICT    260    260       G -> A (IN REF. 1).
FT  CONFLICT    325    325       P -> A (IN REF. 1).
SQ  SEQUENCE   559 AA;  63122 MW;  8CCEE2BDB94514D9 CRC64;
```

```
Query Match             77.7%;  Score 1601; DB 1;  Length 559;
Best Local Similarity   79.5%;  Pred. No. 2e-131;
Matches   283;  Conservative   34;  Mismatches   37;  Indels    2;  Gaps    2;
```

```
Qy     21 AASEGNS-DCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHN 79
             : | |||| |  ||||  ||| |||||||:||  |||  :|||||||| ||||||||:||
Db    204 ACPKGKSEDCYVGKGVTYRGTHSLTTSQASCLPWNSIVLMGKSYTAWRTNSQALGLGRHN 263

Qy     80 YCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQ 139
          |||||||||:|||||  ||:|||||||| |||||||||||:|  |||||| || ||||||
Db    264 YCRNPDGDARPWCHVMKDRKLTWEYCDMSPCSTCGLRQYKRPQFRIKGGLYTDITSHPWQ 323

Qy    140 AAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERRPPHHLTVILGRTYRVVPGEEEQ 199
          ||:| ||:|||||||||||| |||||| |||||||:|| |||||| |||||||||||||||
Db    324 APIFVKNKRSPGERFLCGGVLISSCWVLSAAHCFLERPPNHLKVVLGRTYRVVPGEEEQ 383

Qy    200 KFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECE 259
          || :||| ||||||||||||||||||||| : ||||||||  :|||| ||:|||||||||
Db    384 TFEIEKYIVHEEFDDDTYDNDIALLQLRSQSKQCAQESSSVGTACLPDNLQLPDWTECE 443

Qy    260 LSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANL 319
          |||||||| |||| | ||||||||||||||||||| ::||||| ||||||||||||| :|
Db    444 LSGYGKHEASSPFFSDRLKEAHVRLYPSSRCTSQHLFNKTVTNNMLCAGDTRSGGNQ-DL 502
```

DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; Signal; Multigene family.
FT   SIGNAL   1   36   POTENTIAL.
FT   CHAIN   37   394   SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT   DOMAIN   45   126   KRINGLE.
FT   DOMAIN   142   394   SERINE PROTEASE.
FT   ACT_SITE   189   189   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   238   238   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   345   345   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID   45   126   BY SIMILARITY.
FT   DISULFID   66   108   BY SIMILARITY.
FT   DISULFID   97   121   BY SIMILARITY.
FT   DISULFID   131   262   BY SIMILARITY.
FT   DISULFID   174   190   BY SIMILARITY.
FT   DISULFID   182   251   BY SIMILARITY.
FT   DISULFID   276   351   BY SIMILARITY.
FT   DISULFID   308   324   BY SIMILARITY.
FT   DISULFID   341   369   BY SIMILARITY.
FT   CARBOHYD   315   315   N-LINKED (GLCNAC...). (POTENTIAL).
SQ   SEQUENCE   394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match   70.4%; Score 1451.5; DB 1; Length 394;
Best Local Similarity   67.6%; Pred. No. 1.3e-118;
Matches 263; Conservative 42; Mismatches 69; Indels 15; Gaps 2;

Qy   3 KTAIAIAVALAG----------FATVAQAASEGNSDCYFGNGSAYRGTHSLTESGAS   49
            ||||| : : : ||| : : : : : : ||||||:|||||||
Db   6 KTKLLCVLLCGAVFSLPRQETYRQLARGSRAYGDPHATCYKDQGVTYRGTWSTSESGAQ   65

Qy   50 CLPWNSMLIGKVYTAQNPSAQALGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPS 109
            |:||| ||| ||| : | : ||| ||||||||||| :||||:: : : |||||||:|
Db   66 CINWNSNLLIRRTYNGRMPEAVKLGLGNHNYCRNPDGASKPWCYVIKARKFTSESCSVPV 125

Qy   110 CS--TCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKGHRRSPGERFLCGGILISSCWIL 167
            || ||||||| |||| |||| |||||||||||||| ||:||||||||||||||| :|||
Db   126 CSKATCGLRKYKEPQLHSTGGLFTDITSHPWQAAIFAQNRRSSGERFLCGGILISSCWVL 185

Qy   168 SAAHCFQERFPPHHLTVILGRTYRVVPGEEEOKFEVEKYIVHKEFDDDTYNDIALLQLK 227
            :||||||:| |: || | |: | |||||||| ||| |:|:|||| |||| |||||||||
Db   186 TAAHCFQERYPPQHLRVVLGRTYRVRVKPGKEEQTFEVEKCIVHEEFDDDTYNNDIALLQLK 245

Qy   228 SDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPS 287
            : |: |||| |:: :||| ||||:|||||||||||||||| :|||||: |:||| |||||
Db   246 SGSPQCAQESDSVRAICLPEANLQLPDWTECELSGYGVGKHKSSSPFYSEQLKEGHVRLYPS 305

Qy   288 SRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISW 347
            ||||| ||:| ||| ||||||||||||| |||:||||||||||||| |:||||||||||||
Db   306 SRCTSKFLFNKTVTNNMLCAGDTRSGEIYPNVHDACQGDSGGPLVCMNDNHMTLLGIISW 365

Qy   348 GLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
            |||||| ||| |||||||||||||||||||
Db   366 GVGCGEKDIPGVYTKVTNYLGWIRDNMRP 394

RESULT 6
URT2_DESRO
ID   URT2_DESRO   STANDARD;   PRT;   477 AA.
AC   P15638;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE   alpha-2) (BAT-PA) (T-plasminogen activator).
OS   Desmodus rotundus (Vampire bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC   Desmodontinae; Desmodus.
OX   NCBI_TaxID=9430;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=92039036; PubMed=1937019;

RA   Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA   Alagon A., Donner P., Schleuning W.D.;
RT   "The plasminogen activator family from the salivary gland of the
RT   vampire bat Desmodus rotundus: cloning and expression.";
RL   Gene 105:229-237(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Salivary gland;
RX   MEDLINE=90036867; PubMed=2509450;
RA   Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA   Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT   "Isolation, characterization, and cDNA cloning of a vampire bat
RT   salivary plasminogen activator";
RL   J. Biol. Chem. 264:17947-17952(1989).
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=93331059; PubMed=1309059;
RA   Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA   Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA   Donner P.;
RT   "Plasminogen activators from the saliva of Desmodus rotundus (common
RT   vampire bat): unique fibrin specificity.";
RL   Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC   -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC   EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC   AGENT.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC   plasminogen to form plasmin.
CC   -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC   THE PRESENCE OF FIBRIN I.
CC   -!- SUBUNIT: Monomer.
CC   -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC   AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC   STIMULATION OF ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M63988; AAA31593.1; -.
DR   EMBL; J05082; AAA31596.1; -.
DR   PIR; JS0598; JS0598.
DR   HSSP; P98119; 1A5I.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctn1.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.

```
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE   beta).
OS   Desmodus rotundus (Vampire bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC   Desmodontinae; Desmodus.
OX   NCBI_TaxID=9430;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=92039036; PubMed=1937019;
RA   Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA   Alagon A., Donner P., Schleuning W.D.;
RT   "The plasminogen activator family from the salivary gland of the
RT   vampire bat Desmodus rotundus: cloning and expression.";
RL   Gene 105:229-237(1991).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=93393059; PubMed=1309059;
RA   Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA   Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA   Donner P.;
RT   "Plasminogen activators from the saliva of Desmodus rotundus (common
RT   vampire bat): unique fibrin specificity.";
RL   Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC   -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC       EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC       AGENT.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SUBUNIT: Monomer.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; M63989; AAA31594.1; -.
DR   PIR; JS0599; JS0599.
DR   HSSP; P98119; 1A5I.
DR   MEROPS; S01.239; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Signal; Multigene family.
```

```
FT   SIGNAL        1     36       POTENTIAL.
FT   CHAIN        37    477       SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
FT   DOMAIN       40     82       FIBRONECTIN TYPE-I.
FT   DOMAIN       83    121       EGF-LIKE.
FT   DOMAIN      128    209       KRINGLE.
FT   DOMAIN      225    477       SERINE PROTEASE.
FT   ACT_SITE    272    272       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    321    321       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    428    428       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     42     72       BY SIMILARITY.
FT   DISULFID     70     79       BY SIMILARITY.
FT   DISULFID     87     98       BY SIMILARITY.
FT   DISULFID     92    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    128    209       BY SIMILARITY.
FT   DISULFID    149    191       BY SIMILARITY.
FT   DISULFID    180    204       BY SIMILARITY.
FT   DISULFID    214    345       BY SIMILARITY.
FT   DISULFID    257    273       BY SIMILARITY.
FT   DISULFID    265    334       BY SIMILARITY.
FT   DISULFID    359    434       BY SIMILARITY.
FT   DISULFID    391    407       BY SIMILARITY.
FT   DISULFID    424    452       BY SIMILARITY.
FT   CARBOHYD    185    185       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    398    398       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT    403    403       N -> K (IN REF. 2).
FT   CONFLICT    417    417       Y -> H (IN REF. 2).
FT   CONFLICT    435    435       M -> R (IN REF. 2).
SQ   SEQUENCE   477 AA;  53719 MW;  17486555C0E5077C CRC64;

Query Match             69.3%;  Score 1428;  DB 1;  Length 477;
Best Local Similarity   70.2%;  Pred. No. 1.8e-116;
Matches   254;  Conservative   40;  Mismatches   66;  Indels   2;  Gaps   1;

Qy     17  TVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLG   76
               ::  :||| || | ::|||||::.|:||::||||::||||: : ||: ||.:||| |
Db    116  TGKQCEVDTHATCYKDQGVTYRGTWSTSESGAQCINWNSNLLTRRTYNGRRSDAITLGLG  175

Qy     77  KHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPQFRIKGGLFADIA  134
               ||||:|||  ::|: :|: ||:|::|| : || :|       :|: :|: |||::|||:
Db    176  NHNYCRNPDNNSKPWCYVIKASKPILEFCSVPVCSKATCGLRKYKEPQLHSTGGLFTDIT  235

Qy    135  SHPWQAAIFAKHRRSPGERFLCGGILLSSCWIISAAHCFQERPPPHHLTVILGRTYRVVP  194
               |||||||||:::||::||||||||||:::||::::|||||||:: ::|::|||:|||||
Db    236  SHPWQAAIFAQNRRSSGERFLCGGILISSCWVLTAAHCFQERYPPQQHLRVVLGRTYRVKP  295

Qy    195  GEEEQXFEVEKYIVHKEPDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPD  254
               :::|.|:||:||  |||:|||||:|||:||||||::::|||:..:|:|.:||||::|||:
Db    296  GKEEQTFEVEKCIVHEEFDDDTYNNDIALLQLKSGSPCAQESDSVRAICLPEANLQLPD   355

Qy    255  WTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGG  314
               ||||||||||::  :||||||||:||||||||||||:||::||::|::||:|||||:|||
Db    356  WTECELSGYGKHKSSSPFYSEQLKEGHVRLYPSSRCTSKFLFNKTVTNNMLCAGDTRSGE  415

Qy    315  PQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNVLDWIRDNM  374
               :||:| ||||||||||:|: :|:|: ||||||||||||:||:||||||||:||||:||||
Db    416  IYPNVHDACQGDSGGPLVCMNDNHMTLLGIISWGVGCGEKDIPGVYTKVTNYLGWIRDNM  475

Qy    375  RP  376

Db    476  RP  477
```

```
RESULT 7
URTB_DESRO
ID   URTB_DESRO    STANDARD;    PRT;    431 AA.
AC   P98121;
```

```
FT   SIGNAL       1     36      POTENTIAL.
FT   CHAIN       37    431      SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT   DOMAIN      37     75      EGF-LIKE.
FT   DOMAIN      82    163      KRINGLE.
FT   DOMAIN     179    431      SERINE PROTEASE.
FT   ACT_SITE   226    226      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   275    275      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   382    382      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    41     52      BY SIMILARITY.
FT   DISULFID    46     63      BY SIMILARITY.
FT   DISULFID    65     74      BY SIMILARITY.
FT   DISULFID    82    163      BY SIMILARITY.
FT   DISULFID   103    145      BY SIMILARITY.
FT   DISULFID   134    158      BY SIMILARITY.
FT   DISULFID   168    299      BY SIMILARITY.
FT   DISULFID   211    227      BY SIMILARITY.
FT   DISULFID   219    288      BY SIMILARITY.
FT   DISULFID   313    388      BY SIMILARITY.
FT   DISULFID   345    361      BY SIMILARITY.
FT   DISULFID   378    406      BY SIMILARITY.
FT   CARBOHYD   139    139      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   352    352      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   SEQUENCE   431 AA;  48221 MW;  699B5E675B162CBF CRC64;

Query Match             69.2%;  Score 1427;  DB 1;  Length 431;
Best Local Similarity   69.9%;  Pred. No. 1.9e-116;
Matches  253;  Conservative   41;  Mismatches   66;  Indels    2;  Gaps    1;

Qy     17 TVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQAIGLG 76
             |    :::|  ||| |: :||||    :||| ||   | :|  | :    :  || |||
Db     70 TGKQCEVDTHATCYKDQGVTYRGTWSTSESGAQCINWNSNLLTRRTYNGRRSDAITLGLG 129

Qy     77 KHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPFRIKGGLFADIA 134
          :|||||||| :|||||:: |  :: |:||||:||| ||||||:|::| |  || |: ||
Db    130 NHNYCRNPDNNSKPWCYVIKASKFILEFCSVPVCSKATCGLRKYKEPQLHSTGGLFTDIT 189

Qy    135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHHLTVILGRTYRVVP 194
          |||||||||:: || ||||||||||||||||: |||||||||  | ||:| :|||||| |
Db    190 SHPWQAAIFAQNRRSSGERFLCGGILISSCWLTAAHCFQERYPPQHLRVVLGRTYRVKP 249

Qy    195 GEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPD 254
          |:||| |||||  :|:|||||||:||||||||||.|  |||||.|||| ||| |:|||||
Db    250 GKEEQTFEVEKCIIHEEFDDDTYNNDIALLQLKSGSPQCAQESDSVRAICLPEANLQLPD 309

Qy    255 WTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGG 314
          ||||||||||||  ::| ||| ||| |||||||||||: | ::|||:|||||||||||| 
Db    310 WTECELSGYGKHKSSSPPYSEQLKEGHVRLYPSSRCTSKFLFNKTVTNNMLCAGDTRSGE 369

Qy    315 PQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVVTKVTNYLDWIRDNM 374
             |:||||||||||||||:|:|  ||:||||||| ||| :||:||||||||| |||||||
Db    370 IYPNVHDACQGDSGGPLVCMNDNHMTLLGIISWGVGCGGEKDIPGVVTKVTNYLGWIRDNM 429

Qy    375 RP 376
          ||
Db    430 RP 431
```

```
RESULT 8
URT1_DESRO
ID   URT1_DESRO     STANDARD;      PRT;    477 AA.
AC   P98119;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE   alpha-1).
OS   Desmodus rotundus (Vampire bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC   Desmodontinae; Desmodus.
OX   NCBI_TaxID=9430;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=92039036; PubMed=1937019;
RA   Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA   Alagon A., Donner P., Schleuning W.D.;
RT   "The plasminogen activator family from the salivary gland of the
RT   vampire bat Desmodus rotundus: cloning and expression.";
RL   Gene 105:229-237(1991).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=93393059; PubMed=1309059;
RA   Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA   Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA   Donner P.;
RT   "Plasminogen activators from the saliva of Desmodus rotundus (common
RT   vampire bat): unique fibrin specificity.";
RL   Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC   TISSUE=Salivary gland;
RX   MEDLINE=98022741; PubMed=9354616;
RA   Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RA   Schleuning W.D., Bode W.;
RT   "Catalytic domain structure of vampire bat plasminogen activator: a
RT   molecular paradigm for proteolysis without activation cleavage.";
RL   Biochemistry 36:13483-13493(1997)
CC   -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC       EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC       THE PRESENCE OF FIBRIN I.
CC   -!- SUBUNIT: Monomer.
CC   -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC       AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC       STIMULATION OF ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M63987; AAA31591.1; -.
DR   EMBL; M63986; AAA31592.1; -.
DR   PIR; JS0597; JS0597.
DR   PDB; 1A5I; 23-MAR-99.
DR   MEROPS; S01.232; -.
DR   GlycoSuiteDB; P98119; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctn1.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
```

```
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
FT   SIGNAL        1     36
FT   CHAIN        37    477       SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.
FT   DOMAIN       40     82       FIBRONECTIN TYPE-I.
FT   DOMAIN       83    121       EGF-LIKE.
FT   DOMAIN      128    209       KRINGLE.
FT   DOMAIN      225    477       SERINE PROTEASE.
FT   ACT_SITE    272    272       CHARGE RELAY SYSTEM.
FT   ACT_SITE    321    321       CHARGE RELAY SYSTEM.
FT   ACT_SITE    428    428       CHARGE RELAY SYSTEM.
FT   DISULFID     42     72       BY SIMILARITY.
FT   DISULFID     70     79       BY SIMILARITY.
FT   DISULFID     87     98       BY SIMILARITY.
FT   DISULFID     92    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    128    209       BY SIMILARITY.
FT   DISULFID    149    191       BY SIMILARITY.
FT   DISULFID    180    204       BY SIMILARITY.
FT   DISULFID    214    345       BY SIMILARITY.
FT   DISULFID    257    273       BY SIMILARITY.
FT   DISULFID    265    334       BY SIMILARITY.
FT   DISULFID    359    434       BY SIMILARITY.
FT   DISULFID    391    407       BY SIMILARITY.
FT   CARBOHYD    424    452       N-LINKED (GLCNAC...  ).
FT                               /FTId=CAR_000027.
FT   CARBOHYD    153    153       N-LINKED (GLCNAC...  ).
FT                               /FTId=CAR_000028.
FT   CARBOHYD    398    398
FT   TURN        214    215
FT   STRAND      223    224
FT   TURN        226    227
FT   STRAND      230    231
FT   STRAND      234    236
FT   HELIX       238    239
FT   TURN        240    245
FT   STRAND      254    263
FT   TURN        264    265
FT   STRAND      266    269
FT   HELIX       271    273
FT   TURN        280    282
FT   STRAND      284    287
FT   TURN        297    298
FT   STRAND      300    309
FT   TURN        311    312
FT   TURN        315    317
FT   TURN        319    320
FT   STRAND      323    328
FT   STRAND      338    340
FT   TURN        339    341
FT   STRAND      345    350
FT   TURN        349    356
FT   TURN        355    363
FT   STRAND      358    366
FT   TURN        366    374
FT   STRAND      374    385
FT   HELIX       379    390
FT   TURN        388    393
FT   TURN        393    398
FT   TURN        403    404
FT   STRAND      405    409
FT   TURN        425    426
FT   TURN        428    429
FT   STRAND      431    436
FT   STRAND      437    438
FT   STRAND      439    448
```

```
FT   TURN         455    456
FT   STRAND       459    463
FT   HELIX        464    467
FT   HELIX        468    474
SQ   SEQUENCE    477 AA;  53616 MW;  AA06FD1733C10E5E CRC64;

Query Match             67.4%;  Score 1389.5; DB 1;  Length 477;
Best Local Similarity   69.2%;  Pred. No. 3.9e-113;
Matches   252;  Conservative  37;  Mismatches  72;  Indels  3;  Gaps  2;

Qy    13  AGFATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQA   72
          |!:|||||||||!|||||||||||||| : ||:|||||||||||||:||| : ||||
Db   113  AGY-TGKRCEVDTRATCYEGQGVTYRGTWSTAESRVECINWNSLLTRRTYNGRMPDAFN   171

Qy    73  LGLGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCS--TCGLRQYSQPQFRIKGGLF   130
          ||||:||||||||:|:|||||:|:|  :|::.|| ::: : :||||||:|::|||:::|||
Db   172  LGLGNHNYCRNPNGAPKPWCYVIKAGKPTSESCSVPVCSKATCGLRKYKEPQLHSTGGLF   231

Qy   131  ADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTY   190
          :||:|||||||||:|.||| :||||||||||||||||||||||||||| :  .:.||||||||
Db   232  TDITSHPWQAAIFAQNRRSSGERFLCGGILISSCWVLTAAHCFQESYLPDQLKVVLGRTY   291

Qy   191  RVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADL   250
          ||.||||||.|.|.|||||||||||||.|||||||||||||:.||||| .||.:|||.|.|
Db   292  RVKPGEEEQTFKVKKYIVHKEFDDDTYNNDIALLQLKSDSPQCAQESDSVRAICLPEANL   351

Qy   251  QLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDT   310
          ||||||||||||||||:::|||||:.||:.||||||||||:  : :|:|||.||||||||
Db   352  QLPDWTECELSGYGKHKSSSPFYSEQLKEGHVRLYPSSRCAPKFLFNKTVTNNMLCAGDT   411

Qy   311  RSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWI   370
          |||  : |:|||||||||||||||:||:::||:||||||:|||:|||||||||||||||:|
Db   412  RSGEIYPNVHDACQGDSGGPLVCMNDNHMTLLGIISWGVGCGEKDVPGVYTKVTNYLGWI   471

Qy   371  RDNM   374
          ||||
Db   472  RDNM   475
```

```
Db  126  PDNQKRPWCYVQIGLRQFVQECMVHDCSLSKKPSSSVDQQGFQCG-QKALRPRRKIVGGE  184
                    : ||:|:||||  :||     : :||||||:| :||||||  |::|||||||
Qy  130  FADIASHPWQAAIFAKHRRSPGERPLCGGILISSCWILSAAHCPQERPPHHLTVILGRT   189

Db  185  FTEVENQPWFAAIYQKNKGGSPPSFKCCGSLISPCWVASAAHCPIQLPKKENYVVYLGQS  244
                   |||||| |:|:||:  :|:    |||  :||||||||  :||||  ||:||||||
Qy  190  YRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVRTVCLPP   247

Db  245  KESSYNPGEMKFEVQLIHEYYREDSLAYHNDIALLKIRTSTGQCAQPSRSIQTICLPP  304
               |:|:  :||:|:|||  |   :|   ||   |::|:  ||||: :|||||||||||
Qy  248  ADIQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA  307

Db  305  RFTDAPFGSDCEITGFGKESESDYLYPKNLKMSVVKLVSHEQCMQPHYYGSEINYKMLCA  364
               |:  :||||||||||||||||||||| ||| |:|:|:||||||  :|:|||:|||::|:
Qy  308  GDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYL  367

Db  365  AD-----PEWKT-DSCKGDSGGPLICNIEGRPTLSGIVSWGRGCABKNKPGVTRVSHFL  418
               |:                |||:::
Qy  368  DWIRDNM  374
               |||:::
Db  419  DWIQSHI  425
```

```
RESULT 10
UROK_HUMAN
ID   UROK_HUMAN      STANDARD;      PRT;    431 AA.
AC   P00749; Q15844; Q16618; Q969W6;
DT   21-JUL-1986 (Rel. 01, Created)
DT   20-MAR-1987 (Rel. 04, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE   (U-plasminogen activator).
GN   PLAU.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85215647; PubMed=2987867;
RA   Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT   "The human urokinase-plasminogen activator gene and its promoter.";
RL   Nucleic Acids Res. 13:2759-2771(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RA   Steffens G.J., Heyneker H.L.;
RT   "Cloning and expression of the gene for pro-urokinase in Escherichia
RT   coli.";
RL   Biotechnology 3:923-929(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86056954; PubMed=2415429;
RA   Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RA   Nishida M., Suyama T.;
RT   "Molecular cloning of cDNA coding for human preprourokinase.";
RL   Gene 36:183-188(1985).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85203359; PubMed=3888571;
RA   Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RA   van Elsen A., Herzog A., Bollen A.;
RT   "Molecular cloning, sequencing, and expression in Escherichia coli of
RT   human preprourokinase cDNA.";
RL   DNA 4:139-146(1985).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=lung;
```

```
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; X02389; CAA26231.1; -.
DR   EMBL; M17922; AAA40539.1; -.
DR   PIR; A29420; UKMS.
DR   HSSP; P00749; 1KDU.
DR   MEROPS; S01.231; -.
DR   MGD; MGI:97611; Plau.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Zymogen; Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    433       UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN        21    178       CHAIN A (BY SIMILARITY).
FT   CHAIN       157    178       SHORT A CHAIN (A1).
FT   CHAIN       180    433       CHAIN B (BY SIMILARITY).
FT   DOMAIN       28     64       EGF-LIKE.
FT   DOMAIN       71    152       KRINGLE.
FT   DOMAIN      153    179       CONNECTING PEPTIDE.
FT   DOMAIN      180    433       SERINE PROTEASE.
FT   DISULFID     32     40       BY SIMILARITY.
FT   DISULFID     34     52       BY SIMILARITY.
FT   DISULFID     54     63       BY SIMILARITY.
FT   DISULFID    169    301       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    211    227       BY SIMILARITY.
FT   DISULFID    219    290       BY SIMILARITY.
FT   DISULFID    315    384       BY SIMILARITY.
FT   DISULFID    347    363       BY SIMILARITY.
FT   DISULFID    374    402       BY SIMILARITY.
FT   ACT_SITE    226    226       CHARGE RELAY SYSTEM.
FT   ACT_SITE    277    277       CHARGE RELAY SYSTEM.
FT   ACT_SITE    378    378       CHARGE RELAY SYSTEM.
SQ   SEQUENCE   433 AA;  48268 MW;  A99C35F6250443F9 CRC64;

Query Match             37.8%;  Score 779.5;  DB 1;  Length 433;
Best Local Similarity   40.3%;  Pred. No. 3.4e-60;
Matches  148;  Conservative  67;  Mismatches 129;  Indels 23;  Gaps  5;

Qy   24  EGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRN  83
             : :||| :||||    || :|| : ||    :: :||    :|||||||||||||||
Db   66  DASKTCYHGNGDSYRGKANTDTKGRPCLAWNAPAVLQKPYNAHRPDAISLGLGKHNYCRN  125

Qy   84  PDGDAKPWCHVLKNRRLTWEYCDVPSCS----------------TCGLRQYSQPFRIKGGL  129
            : :|||:|   | |  |  |  | |               ||  :  :|:|| || ||
```

RT   dimensional NMR.";
RL   Nature 337:579-582(1989).
RN   [14]
RP   STRUCTURE BY NMR OF 67-155.
RX   MEDLINE=93003110; PubMed=1327118;
RA   Li X., Smith R.A.G., Dobson C.M.;
RT   "Sequential 1H NMR assignments and secondary structure of the kringle
RT   domain from urokinase.";
RL   Biochemistry 31:9562-9571(1992).
RN   [15]
RP   STRUCTURE BY NMR OF 67-155.
RX   MEDLINE=94149701; PubMed=8107091;
RA   Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT   "Solution structure of the kringle domain from urokinase-type
RT   plasminogen activator.";
RL   J. Mol. Biol. 235:1548-1559(1994).
RN   [16]
RP   VARIANT LEU-141.
RX   MEDLINE=9618279; PubMed=852631;
RA   Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RA   Sawasaki Y., Hanada K.;
RT   "Characterization of single chain urokinase-type plasminogen
RT   activator with a novel amino-acid substitution in the kringle
RT   structure.";
RL   Biochim. Biophys. Acta 1293:83-89(1996).
RN   [17]
RP   VARIANT LEU-141.
RX   MEDLINE=97218551; PubMed=9065988;
RA   Conne B., Berczy M., Belin D.;
RT   "Detection of polymorphisms in the human urokinase-type plasminogen
RT   activator gene.";
RL   Thromb. Haemost. 77:434-435(1997).
RN   [18]
RP   ERRATUM.
RA   Conne B., Berczy M., Belin D.;
RL   Thromb. Haemost. 78:973-973(1997).
RN   [19]
RP   VARIANT LEU-141.
RX   MEDLINE=97337920; PubMed=9194591;
RA   Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
RA   Creutzburg S., Graeff H., Magdolen V.;
RT   "Mutational analysis of the genes encoding urokinase-type plasminogen
RT   activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
RL   Electrophoresis 18:686-689(1997).
CC   -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
CC   THERAPY OF THROMBOLYTIC DISORDERS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC   plasminogen to form plasmin.
CC   -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC   OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC   LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC   MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
CC   -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC   in Pulmonary Embolism (PE) to initiates fibrinolysis.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X02419; CAA26268.1; -.
DR   EMBL; M15476; AAA61253.1; -.
DR   EMBL; D00244; BAA00175.1; -.
DR   EMBL; D11143; BAA01919.1; -.
DR   EMBL; X02760; CAA26535.1; -.
DR   EMBL; AF377330; AAK53822.1; -.
DR   EMBL; BC013575; AAH13575.1; -.

RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SEQUENCE OF 66-431 FROM N.A.
RX   MEDLINE=84272706; PubMed=6589620;
RA   Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
RT   "Identification and primary sequence of an unspliced human urokinase
RT   poly(A)+ RNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN   [8]
RP   SEQUENCE OF 21-177.
RX   MEDLINE=83055084; PubMed=6754569;
RA   Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RA   Flohe L.;
RT   "The primary structure of high molecular mass urokinase from human
RT   urine. The complete amino acid sequence of the A chain.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN   [9]
RP   SEQUENCE OF 156-176 AND 179-224.
RX   MEDLINE=83003608; PubMed=6749491;
RA   Schaller J., Nick H., Rickli E.E., Gillessen D., Lergier W.,
RA   Studer R.O.;
RT   "Human low-molecular-weight urinary urokinase. Partial
RT   characterization and preliminary sequence data of the two polypeptide
RT   chains.";
RL   Eur. J. Biochem. 125:251-257(1982).
RN   [10]
RP   SEQUENCE OF 158-410.
RX   MEDLINE=83055099; PubMed=6754572;
RA   Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
RT   "The complete amino acid sequence of low molecular mass urokinase
RT   from human urine.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS)
RX   MEDLINE=96000858; PubMed=8591045;
RA   Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA   Dobson C.M., Stuart D.I., Jones E.Y.;
RT   "The crystal structure of the catalytic domain of human
RT   urokinase-type plasminogen activator.";
RL   Structure 3:681-691(1995).
RN   [12]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX   MEDLINE=20266327; PubMed=10805774;
RA   Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA   Bode W., Magdolen V., Huber R., Moroder L.;
RT   "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT   selective inhibitors of human urokinase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN   [13]
RP   STRUCTURE BY NMR.
RX   MEDLINE=89127526; PubMed=2536903;
RA   Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RT   "Dynamics of the multidomain fibrinolytic protein urokinase from two-

(Left margin tag columns for the above blocks:)

RX RA RA RA RA RA RA RA RA RA RA RA RA RA RA RA RA RT RT RL RN RP RX RA RT RT RL RN RP RX RA RA RT RT RL RN RP RX RA RT RT RL RN RP RX RA RT RT RL RN RP RX RA RA RT RT RL RN RP RX RA RT RT RL RN RP RX RA RT RT RL RN RP RX RA RT

DR   EMBL; K03226; AAC97138.1; -.
DR   EMBL; K02286; AAA61252.1; -.
DR   EMBL; A21571; CAA01559.1; -.
DR   EMBL; A18397; CAA01390.1; -.
DR   PIR; A00931; UKHU.
DR   PDB; 1KDU; 31-OCT-93.

Query Match   37.7%; Score 777; DB 1; Length 431;
Best Local Similarity   43.3%; Pred. No. 5.6e-60; -.
Matches 157; Conservative 48; Mismatches 134; Indels 24; Gaps   6;

Qy   29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
              ||||| ||| |||||| ||: || :| |||||||||||
Db   70 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQQTYHAHRSDALQLGLGKHNYCRNPDNRR 129

Qy   89 KPWCHVLKNRRLTWEYCDVPSCS----------------TCGLRQYSQPQFRIKGGLFADIA 134
              :||:|| |:|::|| || |: : |||||||:||: ||||
Db   130 RPWCYVQVGLKPLVQECMVHDCADGKKPSSPPELKFQCG-QKTLRPRFKIIGGEFTTIE 188

Qy   135 SHPWQAAIFAKHRRSPGERFLCGGILLSSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
              :||| |:|||||||||||||| :|:|:|: ||| :||| |||:
Db   189 NQPWFAAIYRRH-RGGSVTYVCGGSLMSPCWVISATHCFIDYPKKBDYIVLGRSRLNSN 247

Qy   195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQL 252
              ||| || ||:| |: ||::||| |||||| :|: :|:
Db   248 TQGEMKFEVENLILHKDYSADTLAHHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDP 307

Qy   253 PDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRS 312
              | :|:::|| :|||| ||| ||||||
Db   308 QFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECQQPHYYGSEVTTKMLCAAD--- 364

Qy   313 GGPQANLHDACQGDSGGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDMWIRD 372
              ||| ||| |||||||||||| :||||||||:|||| |||| |:|||:||:|:
Db   365 --PQWKT-DSCCGDSGGPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFLPWIRS 421

Qy   373 NMR 375
              : :
Db   422 HTK 424

RESULT 11
UROK_PAPCY
ID   UROK_PAPCY   STANDARD;   PRT;   433 AA.
AC   P16227;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE   (U-plasminogen activator).
GN   PLAU.
OS   Papio cynocephalus (Yellow baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thoracic aorta;
RX   MEDLINE=90287734; PubMed=2113276;
RA   Au Y.P.T.; Wang T.W.; Clowes A.W.;
RT   "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT   plasminogen activator.";
RL   Nucleic Acids Res. 18:3411-3411(1990).
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC   plasminogen to form plasmin.
CC   -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC   OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC   LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC   MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.

DR   EMBL; X51935; CAA36200.1; -.
DR   PIR; S14687; UKBAY.
DR   HSSP; P00749; 1LMW.
DR   MEROPS; S01.231; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Zymogen; Signal.
FT   SIGNAL   1   20   POTENTIAL.
FT   CHAIN   21   433   UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN   21   176   CHAIN A (BY SIMILARITY)
FT   CHAIN   155   176   SHORT A CHAIN (A1) (BY SIMILARITY).
FT   CHAIN   178   433   CHAIN B (BY SIMILARITY).
FT   DOMAIN   26   62   EGF-LIKE.
FT   DOMAIN   69   150   KRINGLE.
FT   DOMAIN   151   177   CONNECTING PEPTIDE.
FT   DOMAIN   178   433   SERINE PROTEASE.
FT   DISULFID   30   38   BY SIMILARITY.
FT   DISULFID   32   50   BY SIMILARITY.
FT   DISULFID   52   61   BY SIMILARITY.
FT   DISULFID   167   298   INTERCHAIN (BY SIMILARITY).
FT   DISULFID   208   224   BY SIMILARITY.
FT   DISULFID   216   287   BY SIMILARITY.
FT   DISULFID   315   384   BY SIMILARITY.
FT   DISULFID   347   363   BY SIMILARITY.
FT   DISULFID   374   402   BY SIMILARITY.
FT   ACT_SITE   223   223   CHARGE RELAY SYSTEM.
FT   ACT_SITE   274   274   CHARGE RELAY SYSTEM.
FT   ACT_SITE   378   378   CHARGE RELAY SYSTEM.
FT   CARBOHYD   324   324   N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ   SEQUENCE   433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match   37.6%; Score 774.5; DB 1; Length 433;
Best Local Similarity   43.5%; Pred. No. 9.2e-60; -.
Matches 160; Conservative 45; Mismatches 132; Indels 31; Gaps   8;

Qy   29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
              ||||| ||| |||||| ||: || :| |||||||||||
Db   69 CYEGNGHFYRGKASTDTMGRSCLAWNSATVLQQTYHAHRSDALQLGLGKHNYCRNPDNRR 128

Qy   89 KPWCHVLKNRRLTWEYCDVPSCS----------------TCGLRQYSQPQFRIKGGLFADIA 134
              :||:|| |:|::|| || |: : |||||||:||: ||||
Db   129 RPWCYVQGLKQRVQECMVHNCADGKKPSSPPELQFCGQRTL-RPRRKIVGGEFTTIE 187

Qy   135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
              :||| |:|||||||| ||||:|:|:|:|:
Db   188 NQPWFAAIYRRH-RGGSVTYVCCGSLISPCWVVSATHCFINYPKKBDYIVYLGRSRLNSN 246

```
Qy    195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLP-----P 247
              :  |||||| |:|::|  |  ||||  | :|:|:|   |:|| :|
Db    247 TQGEMKFEVENLIHEDYSADTLAHHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDP 306

Qy    248 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 307
              ||||  | ||:||  |  :|:|| : |:| :| :||||||||||  ||||
Db    307 ND--PPFGTSCEITGFGKENSTDYLYPEQLKMTVKLVSHQKCQOPHYYGSEVTTKMLCA 364

Qy    308 GDTRSGGPQQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYL 367
              | ||||||||||||  |||||||||||||||||  ||| |||||||:|: :|
Db    365 AD------PQWET-DSCQGDSGGPLVCSIQGHMTLTGIVSWGRGCALKDKPGVTRVSRFL 418

Qy    368 DWIRDNMR 375
              || ||
Db    419 PWIHSHTR 426
```

```
RESULT 12
UROK_RAT
ID    UROK_RAT                STANDARD;      PRT;     432 AA.
AC    P29598;
DT    01-APR-1993 (Rel. 25, Created)
DT    01-APR-1993 (Rel. 25, Last sequence update)
DT    28-FEB-2003 (Rel. 41, Last annotation update)
DE    Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE    (U-plasminogen activator).
GN    PLAU.
OS    Rattus norvegicus (Rat).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX    NCBI_TaxID=10116;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=Fischer 344;
RX    MEDLINE=92233409; PubMed=1568219;
RA    Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA    Kefford R.F.;
RT    "Transcriptional and posttranscriptional activation of urokinase
RT    plasminogen activator gene expression in metastatic tumor cells.";
RL    Cancer Res. 52:2489-2496(1992).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Kidney;
RA    Rabbani S.A.;
RL    Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
CC    -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC        plasminogen to form plasmin.
CC    -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
CC        CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
CC        CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
CC        LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC    -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC    -!- SIMILARITY: Contains 1 kringle domain.
CC    -!- SIMILARITY: Contains 1 EGF-like domain.
CC    --------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------
DR    EMBL; X63434; CAA45028.1; -.
DR    EMBL; X65651; CAA46601.1; -.
DR    PIR; S24604; S18932.
DR    HSSP; P00749; 1KDU.
DR    MEROPS; S01.231; -.
DR    InterPro; IPR001314; Chymotrypsin.
DR    InterPro; IPR006209; EGF_like.
DR    InterPro; IPR006210; IEGF.
DR    InterPro; IPR000001; Kringle.
DR    InterPro; IPR001254; Ser_protease_Try.
DR    Pfam; PF00051; kringle; 1.
DR    Pfam; PF00089; trypsin; 1.
DR    PRINTS; PR00722; CHYMOTRYPSIN.
DR    PRINTS; PR00018; KRINGLE.
DR    ProDom; PD000395; Kringle; 1.
DR    SMART; SM00181; EGF; 1.
DR    SMART; SM00130; KR; 1.
DR    SMART; SM00020; Tryp_SPc; 1.
DR    PROSITE; PS00022; EGF_1; 1.
DR    PROSITE; PS01186; EGF_2; FALSE_NEG.
DR    PROSITE; PS00021; KRINGLE_1; 1.
DR    PROSITE; PS50070; KRINGLE_2; 1.
DR    PROSITE; PS00240; TRYPSIN_DOM; 1.
DR    PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR    PROSITE; PS00135; TRYPSIN_SER; 1.
KW    Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW    Kringle; EGF-like domain; Zymogen; Signal.
FT    SIGNAL        1     19       POTENTIAL.
FT    CHAIN        20    432       UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT    CHAIN        20    177       CHAIN A (BY SIMILARITY).
FT    CHAIN       156    177       SHORT A CHAIN (A1) (BY SIMILARITY).
FT    CHAIN       179    432       CHAIN B (BY SIMILARITY).
FT    DOMAIN       27     63       EGF-LIKE.
FT    DOMAIN       70    151       KRINGLE.
FT    DOMAIN      152    178       CONNECTING PEPTIDE.
FT    DOMAIN      179    432       SERINE PROTEASE.
FT    DISULFID     31     39       BY SIMILARITY.
FT    DISULFID     33     51       BY SIMILARITY.
FT    DISULFID     53     62       BY SIMILARITY.
FT    DISULFID    168    300       INTERCHAIN (BY SIMILARITY).
FT    DISULFID    210    226       BY SIMILARITY.
FT    DISULFID    218    289       BY SIMILARITY.
FT    DISULFID    314    383       BY SIMILARITY.
FT    DISULFID    346    362       BY SIMILARITY.
FT    DISULFID    373    401       BY SIMILARITY.
FT    ACT_SITE    225    225       CHARGE RELAY SYSTEM.
FT    ACT_SITE    276    276       CHARGE RELAY SYSTEM.
FT    ACT_SITE    377    377       CHARGE RELAY SYSTEM.
FT    CONFLICT     16     16       N -> H (IN REF. 2).
FT    CONFLICT     24     24       E -> G (IN REF. 2).
FT    CONFLICT    332    332       D -> N (IN REF. 2).
SQ    SEQUENCE   432 AA;  47957 MW;  4EB1B9C71G244C8 CRC64;
```

```
Query Match             37.5%;  Score 772.5;  DB 1;  Length 432;
Best Local Similarity   40.6%;  Pred. No. 1.4e-59;
Matches 147;  Conservative 64;  Mismatches 128;  Indels 23;  Gaps 5;
```

```
Qy     29 CYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
              ||: |||  :|||||  | |:| ||||  :| | :::|| | :||||||||||||||||
Db     70 CYHGNGQSYRGKANTDTKGRPCLAWNSPAVLQQTYNAHRSDALSLGLGKHNYCRNPDNQR 129

Qy     89 KPWCHVLKNRRLTWEYCDVPSCS----------------TCGLRQYSQPFRIKGGLFADIA 134
              ||| ::  ||  :|| |::|::       :::|| ::||||
Db    130 RPWCYVQIGLKQFVQECMVQDCSLSKKPSSTVDQQGFQCG-QXALRPRFKIVGGEFTTVE 188

Qy    135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
              :  ||:|| ||||  :||  | :|| ||:  :|||||| |: |||:::|:|||:|
Db    189 NQPWFAAIYLKNKGGSPSFKCGGSLISPCWVASATHCFVNQPKKEBVVVYLGQSKRNSY 248

Qy    195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQL 252
              |||||| |:|::|  ||||  |: |:|:|   |:|| :|
Db    249 NPGEMKFEVEQLIHEDFSDETLAFHNDIALLKIRTSTGCAQPSRTIQTICLPPRFGDA 308

Qy    253 PDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRS 312
              :  :|||   | ||:||||||||| :|:  :|| || ||||| |::|:||:|||||
Db    309 PFGSDCEITGFGQESATDYFYPKDLKMSVVKIISHEQCKQPHYYGSEINYKMLCAAD--- 365

Qy    313 GGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRD 372
              |||:||||||  |||||||||  : ||||||||||||| :|||||||  | : ::|||:
Db    366 --PEWKT-DSCSGDSGGPLICNIDGRPTLSGIVSWGSGCAEKNKPGVVTRVSYFLNWIQS 422

Qy    373 NM 374
```

```
Db          423 HI 424
             ::

RESULT 13
UROK_PIG
ID   UROK_PIG      STANDARD;      PRT;      442 AA.
AC   P04185;
DT   20-MAR-1987 (Rel. 04, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE   (U-plasminogen activator).
GN   PLAU.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=85087954; PubMed=6096832;
RA   Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT   "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL   Nucleic Acids Res. 12:9525-9541(1984).
RN   [2]
RP   REVISION TO 241.
RA   Nagamine Y.;
RL   Submitted (DEC-1986) to the PIR data bank.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   --------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------
DR   EMBL; X01648; CAA25806.1; -.
DR   EMBL; X02724; CAA26511.1; -.
DR   PIR; A00932; UKPG.
DR   HSSP; P00749; 1KDU.
DR   MEROPS; S01.231; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Zymogen; Signal.
FT   SIGNAL         1     20       BY SIMILARITY.
FT   CHAIN         21    442       UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN         21    188       CHAIN A (BY SIMILARITY).
FT   CHAIN        190    442       CHAIN B (BY SIMILARITY).
FT   DOMAIN        29     65       EGF-LIKE.
FT   DOMAIN        72    153       KRINGLE.
FT   DOMAIN       154    189       CONNECTING PEPTIDE.
FT   DOMAIN       190    442       SERINE PROTEASE.
FT   CARBOHYD     152    152       N-LINKED (GLCNAC...).
FT   DISULFID      33     41       BY SIMILARITY.
FT   DISULFID      35     53       BY SIMILARITY.
FT   DISULFID      55     64       BY SIMILARITY.
FT   DISULFID     179    310       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     220    236       BY SIMILARITY.
FT   DISULFID     228    299       BY SIMILARITY.
FT   DISULFID     324    393       BY SIMILARITY.
FT   DISULFID     356    372       BY SIMILARITY.
FT   DISULFID     383    411       BY SIMILARITY.
FT   ACT_SITE     235    235       CHARGE RELAY SYSTEM.
FT   ACT_SITE     286    286       CHARGE RELAY SYSTEM.
FT   ACT_SITE     387    387       CHARGE RELAY SYSTEM.
FT   CONFLICT     241    241       Q -> H (IN REF. 1; CAA25806).
FT   CONFLICT     242    242       Q -> H (IN REF. 1; CAA26511).
FT   CONFLICT     288    288       A -> GS (IN REF. 1; CAA25806).
SQ   SEQUENCE   442 AA;  49116 MW;  EE32FCEF501321EB CRC64;

Query Match    37.3%;   Score 769.5;   DB 1;   Length 442;
Best Local Similarity   41.2%;   Pred. No. 2.6e-59;
Matches  153;   Conservative  50;   Mismatches 135;   Indels  33;   Gaps  6;

Qy    29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
          :|:||||| ||||||| |||| ||||| |||||||||||:|||||||||||||||||||
Db    72 CFEGNGHSYRGKANTNTGGRPCLPWNSATVLLNTYHAHRPDALQLGLGKHNYCRNPDNQR 131

Qy    89 KPWCHVLKNRRLTWEYCDVPSCS---------------------------TCGLRQYSQPQFRI 125
          :||:|               |||:|                             |: :|:|::|
Db   132 RPWCYVQVGLKQLVQECMVPNCSGGESHRPAYDGKNPPSTPEKVEFQCG-QKALRPRFKI 190

Qy   126 KGGLFADIASHPWQAAIPAKHRRSPGERFLCGGILISSCWILSAAHCFOERFPPHHLTVI 185
          ||||||| :|:|| | |   |:||| |:: |||  ||||:|||||:|:||  :||::||||
Db   191 VGGKSTTIENQPWFAAIYRRH-RGGSVTYVCGGSLISPCWVVSATHCFINYQQKEDYIVY 249

Qy   186 LGRTYRVVPGEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTV 243
          ||| | |||||||:||||:|||:||||| :   |||||||||||||:|||::|||::|:
Db   250 LGRQTLHSSTHGEMKFEVEKLILHEDYSADSLAHHNDIALLKIRTDKGGQCAQPSRSIQTI 309

Qy   244 CLPPADLQLPDWTECELSGYGKHEALSPPYSERLKEAHVRLYPSSRCTSQHLLNRTVTDN 303
          ||| : | |||||| |:|:|||:||||| : :| |:|| |||||:||||| : ||:
Db   310 CLPPVNGDAHFGASCEIVGFGKEDPSDYLYPEQLKMTVVKLVSHRECQPHYYGSEVTTK 369

Qy   304 MLCAGDTRSGGPQOANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCCQGKDVPGVYTKV 363
          |||| |    | ||: |  |||:|||||| | |:|||||||||||||:|| :|||| |
Db   370 MLCAAD-----PQWKT-DSCQGDSGGPLVCSTQGRLTLTGIVSWGRECAMKDKPGVTRV 423

Qy   364 TNYLDWIRDNM 374
          : :| ||:|:|
Db   424 SRPLTWIHTHV 434

RESULT 14
UROK_CHICK
ID   UROK_CHICK    STANDARD;      PRT;      434 AA.
AC   P15120;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE   (U-plasminogen activator).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90110185; PubMed=2295632;
RA   Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT   "The chicken urokinase-type plasminogen activator gene.";
```

```
RL   J. Biol. Chem. 265:1339-1344(1990).
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; J05187; AAA49131.1; -.
DR   EMBL; J05188; AAA49130.1; -.
DR   PIR; A35005; A35005.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.231; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF-like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Signal; Zymogen.
FT   SIGNAL        1     20
FT   CHAIN        21    434       UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN        21    171       CHAIN A (BY SIMILARITY).
FT   CHAIN       173    434       CHAIN B (BY SIMILARITY).
FT   DOMAIN       36     72       EGF-LIKE.
FT   DOMAIN       79    158       KRINGLE.
FT   DOMAIN      159    172       CONNECTING PEPTIDE.
FT   DOMAIN      173    434       SERINE PROTEASE.
FT   DISULFID     40     48       BY SIMILARITY.
FT   DISULFID     42     60       BY SIMILARITY.
FT   DISULFID     62     71       BY SIMILARITY.
FT   DISULFID    162    296       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    202    218       BY SIMILARITY.
FT   DISULFID    210    285       BY SIMILARITY.
FT   DISULFID    310    379       BY SIMILARITY.
FT   DISULFID    342    358       BY SIMILARITY.
FT   DISULFID    369    397       BY SIMILARITY.
FT   ACT_SITE    217    217       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    272    272       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    373    373       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD    228    228       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   434 AA;  49400 MW;  BD8811048DD666A55 CRC64;

Query Match             37.0%;  Score 762;  DB 1;  Length 434;
Best Local Similarity   45.4%;  Pred. No. 1.1e-58;
Matches   163;  Conservative  56;  Mismatches 116;  Indels  24;  Gaps  12;

Qy   26  NSDCYFGNGSAYRGTHSLTESGASCLPWN--SMILIGKVYTAQNPSAQALGLGKHNYCRN 83
             || ||||| ||| |: :|.:|::.|  :|   |||||||||||
Db   76  NSICYSGNGEDYRG---MAEDPGCLYWDHPSVIRWGD-YHADLKNALQLGLGKHNYCRN 130
```

```
Qy   84  PDGDAKPWCHVLKNRRLTWEYCD-VPSCS-TCGLRQYSQPQFRIKGGLFADIASHPWQAA 141
             |:.::|||:..|. :::||| |..|| :.:|||. |::.:||.|
Db  131  PNGRSRPWCYTKRRYSIQETPCSTIEKCERTCGQRSFSK-YFKIVGGSQAEVETQPWIAG 189

Qy  142  IFAKHRRSPGERFLCGGILISSCWILSAAHCFQE--RFPPHH--LTVILGRTYRVVPGSE 197
             |.|.::::||. |.:|||.:.:.:||:..||.:|    :.|   |.::|.:|.:..:
Db  190  IFQNIMGT--DQFLCGGSLIDPCWVLTAAHCFNPTKKQPNKSVYKVFIGKSILNTNDEH 247

Qy  198  EQKFEVEKYIVHKEFDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDW 255
             |:.:.:: ||.:.|..: ||||||||||||:|::..||.|.|::|||||||.::|.|
Db  248  EQVFMVDEIISHPDFTDHTGGNDNDIALIRIRTASGQCAVESNYVRTVCLPEKNLNLYDN 307

Qy  256  TECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGP 315
             ||:.|:||:|:||.  :::.||.:|.|.||. ||||:|||||||||||:||||..||||
Db  308  TWCEIAGYGKQNSYDIYAQRLMSATVNLISQDDCKNKYYDSTRVTDNMVCAGD----P 362

Qy  316  QANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNM 374
             ||.|.||.|||||||:||.|||.||||||:|||  ||||.|.|||:||||||:|||:|||
Db  363  LWET-DACKGDSGPMVCEHNGRMTLYGIVSWGDGCAKKNKPGVYTRVYRYLNWIDSNM 420
```

```
RESULT 15
UROK_BOVIN
ID   UROK_BOVIN     STANDARD;      PRT;    433 AA.
AC   Q05589; Q28209;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE   (U-plasminogen activator).
GN   PLAU.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic endothelium;
RX   MEDLINE=93216119; PubMed=8385052;
RA   Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA   Schleuning W.-D.;
RT   "Bovine urokinase-type plasminogen activator and its receptor:
RT   cloning and induction by retinoic acid.";
RL   Gene 125:177-183(1993).
RN   [2]
RP   SEQUENCE OF 12-433 FROM N.A.
RC   TISSUE=Kidney;
RA   Ravn P., Berglund L., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen activators uPA
RT   and tPA.";
RL   Int. Dairy J. 5:605-617(1995).
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC       plasminogen to form plasmin.
CC   -!- INDUCTION: By retinoic acid.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: Contains 1 kringle domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L03546; AAA51419.1; -.
DR   EMBL; X85801; CAA59796.1; -.
DR   PIR; JN0560; JN0560.
DR   HSSP; P00749; 1LMW.
DR   MEROPS; S01.231; -.
DR   InterPro; IPR001314; Chymotrypsin.
```

```
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW   Kringle; EGF-like domain; Signal; Zymogen.
FT   SIGNAL        1     20
FT   CHAIN        21    433       UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT   CHAIN        21    179       CHAIN A (BY SIMILARITY).
FT   CHAIN       181    433       CHAIN B (BY SIMILARITY).
FT   DOMAIN       29     65       EGF-LIKE.
FT   DOMAIN       72    153       KRINGLE.
FT   DOMAIN      154    180       CONNECTING PEPTIDE.
FT   DOMAIN      181    433       SERINE PROTEASE.
FT   DISULFID     33     41       BY SIMILARITY.
FT   DISULFID     35     53       BY SIMILARITY.
FT   DISULFID     55     64       BY SIMILARITY.
FT   DISULFID    170    301       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    211    227       BY SIMILARITY.
FT   DISULFID    219    290       BY SIMILARITY.
FT   DISULFID    315    384       BY SIMILARITY.
FT   DISULFID    347    363       BY SIMILARITY.
FT   DISULFID    374    402       BY SIMILARITY.
FT   ACT_SITE    226    226       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    277    277       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    378    378       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CONFLICT    189    189       A -> T (IN REF. 2).
SQ   SEQUENCE   433 AA;  48730 MW;  4DE1B8D4DA47027A CRC64;

Query Match             36.1%;  Score  745;  DB 1;  Length 433;
Best Local Similarity   41.3%;  Pred. No. 3.4e-57;
Matches  150;  Conservative   54;  Mismatches  135;  Indels   24;  Gaps    6;

Qy         29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
              || |||.:|||| : |:|| :. ::  :|:| :|: .:|||||||||:|||||||||
Db         72 CYQGNGHSYRGKANRDLSGRPCLAWDSPTVLLKMYHAHRSDAIQLGLGKHNYCRNPDNQR 131

Qy         89 KPWCHVLKNRRLTWEYCDVPSCST----------CGLRQYSQPQFRIKGGLFADIA 134
              :||| |: . :|:|| |||||. | ||  ::  :|:|:||
Db        132 RPWCYVQIGLKQFVQFCMVQDCSVGKSPSSPREKEEFQCG-QKALRPRFKIVGGQVTNAE 190

Qy        135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
              . |:|||::.:| :| |||||||| ||| :: :||| |: :  ||
Db        191 NQPWFAAIYRRH-RGGSITYLCGGSLISPCWVVSATHCFIDHPKKENYIVYLGQSRINSD 249

Qy        195 GEEEQKFEVEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQL 252
              |::|||: : |: : |: |||||| ||| :||||||||
Db        250 TRGEMQFEVEKLILHEDYSAESLAHHNDIALLKIRTSRGQCAQPSRSIQTICLPPEHEDA 309

Qy        253 PDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRS 312
              |||:|| :||| : :||: || :: |||  ||||||||
Db        310 HSRTRCEITGFGKENPSDYRYSDELKMTFVSLVSHEVCQQPHYYGAEVTDRMLCAAD--- 366

Qy        313 GGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRD 372
              |||:|| |:||||||||||||  |||||:|| |||| ||::|:||||||||:| :|:|
Db        367 --PQWET-DSCQGDSGGPLVCTIQGRLTLTGIVSWGRDCAMKYKPGVYTRVSKFLPWINT 423

Qy        373 NMR 375
Db        424 HTR 426
```

```
                Copyright   GenCore version 5.1.6
                          (c) 1993 - 2003   Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August  8, 2003, 17:11:32 ; Search time 39 seconds
                                        (without alignments)
                                        2494.509 Million cell updates/sec

Title:       US-09-987-455-8
Perfect score: 2061
Sequence:    1 MKKTAIAIAVALAGFATVAQ.........GVTKVTNYLDWIRDNMRPG 377

Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched:    830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:         830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_23:*
             1:  sp_archea:*
             2:  sp_bacteria:*
             3:  sp_fungi:*
             4:  sp_human:*
             5:  sp_invertebrate:*
             6:  sp_mammal:*
             7:  sp_mhc:*
             8:  sp_organelle:*
             9:  sp_phage:*
            10:  sp_plant:*
            11:  sp_rodent:*
            12:  sp_virus:*
            13:  sp_vertebrate:*
            14:  sp_unclassified:*
            15:  sp_rvirus:*
            16:  sp_bacteriap:*
            17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1961	95.1	516	4	Q9BU99	Q9bu99 homo sapien
2	1942	94.2	395	4	Q8BZW1	Q9bzw1 homo sapien
3	1669	81.0	562	6	Q8SQ23	Q8sq23 sus scrofa
4	1601	77.7	559	11	Q91VP2	Q91vp2 mus musculu
5	1536	74.5	564	6	Q8MKB1	Q8mkb1 oryctolagus
6	769	37.3	433	6	Q8MIL0	Q8milo oryctolagus
7	767	37.2	433	6	Q8MHY7	Q8mhy7 oryctolagus
8	763	37.0	202	13	Q90675	Q90675 gallus gall
9	742.5	36.0	653	11	Q8VCS4	Q8vcs4 mus musculu
10	709.5	34.4	560	4	Q14520	Q14520 homo sapien
11	698.5	33.9	517	11	Q8K0D2	Q8k0d2 mus musculu
12	667.5	32.4	615	4	Q8IZZ5	Q8izz5 homo sapien
13	634.5	30.8	597	11	O35727	O35727 mus musculu
14	630	30.6	616	6	O97507	O97507 sus scrofa
15	607	29.5	812	11	Q91WJ5	Q91wj5 mus musculu
16	593.5	28.8	812	11	Q9R0W3	Q9r0w3 rattus norv

17	580	28.1	806	6	O18783	O18783 macropus eu
18	560.5	27.2	810	4	Q15146	Q15146 homo sapien
19	555.5	26.6	334	6	O46507	O46507 papio hamad
20	549	26.6	454	6	O46506	O46506 papio hamad
21	525.5	25.5	429	13	Q8AVB0	Q8avb0 brachydanio
22	487.5	23.7	300	4	Q96EF3	Q96ef3 homo sapien
23	463.5	22.5	385	5	Q25101	Q25101 herdmania m
24	453	22.0	855	11	Q9JJ17	Q9jj17 rattus norv
25	452	21.9	761	11	Q99JC8	Q99jc8 rattus norv
26	449	21.8	422	4	Q8WVC1	Q8wvc1 homo sapien
27	443.5	21.5	90	4	Q8NG20	Q8ng20 homo sapien
28	443.5	21.5	802	4	Q8IUE2	Q8iue2 homo sapien
29	443.5	21.5	811	4	Q8IU80	Q8iu80 homo sapien
30	442.5	21.5	868	5	Q9YIV3	Q9yiv3 polyandroca
31	433.5	21.0	608	13	Q9PTW7	Q9ptw7 struthio ca
32	427.5	20.7	799	11	Q9DBI0	Q9dbi0 mus musculu
33	425.5	20.6	214	6	Q9XT70	Q9xt70 oryctolagus
34	423.5	20.5	607	13	Q91001	Q91001 gallus gall
35	422.5	20.5	267	5	Q9BK47	Q9bk47 luidia foli
36	422	20.5	371	11	Q8CJ16	Q8cj16 rattus norv
37	422	20.5	445	11	Q8CJ17	Q8cj17 rattus norv
38	416.5	20.2	471	11	Q8CFE0	Q8cfe0 mus musculu
39	415.5	20.2	505	5	Q966V4	Q966v4 halocynthia
40	409.5	19.9	537	4	Q9BYE1	Q9bye1 homo sapien
41	409.5	19.9	581	4	Q9BYE2	Q9bye2 homo sapien
42	408	19.8	455	11	Q8CDR0	Q8cdr0 mus musculu
43	407.5	19.8	394	5	P91817	P91817 tachypleus
44	405	19.7	613	13	Q03711	Q03711 xenopus lae
45	401	19.5	453	11	Q8VDE0	Q8vde0 mus musculu

ALIGNMENTS

```
RESULT 1
Q9BU99
ID   Q9BU99         PRELIMINARY;        PRT;     516 AA.
AC   Q9BU99;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to plasminogen activator, tissue.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR   EMBL; BC002795; AAH02795.1; -.
DR   HSSP; P00750; 1A5H.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS00070; KRINGLE_2; 2.
```

DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease.
SQ   SEQUENCE   516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match   95.1%; Score 1961; DB 4; Length 516;
Best Local Similarity   99.7%; Pred. No. 5.4e-187;
Matches 355; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy         21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 80
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNY 220

Qy         81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 280

Qy        141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 340

Qy        201 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 FEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 400

Qy        261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 460

Qy        321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 516
```

RESULT 2
Q9BZW1
ID   Q9BZW1   PRELIMINARY;   PRT;   395 AA.
AC   Q9BZW1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Neonatal thrombolytic agent alpha-form (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dou D.;
RT   "A brain-type plasminogen activator.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; AF260825; AAK11956.1; -.
DR   HSSP; P00750; 1PK2.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000083; Fibrnctn1.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD003395; Kringle; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT   NON_TER   395   395
SQ   SEQUENCE   395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match   94.2%; Score 1942; DB 4; Length 395;
Best Local Similarity   99.2%; Pred. No. 3e-185;
Matches 351; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy         23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 101

Qy         83 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 161

Qy        143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKPE 202
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 FAKHEESPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKPE 221

Qy        203 VEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 262
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CRKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSG 281

Qy        263 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 YGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA 341

Qy        323 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 395
```

RESULT 3
Q8SQ23
ID   Q8SQ23   PRELIMINARY;   PRT;   562 AA.
AC   Q8SQ23;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   T-plasminogen activator.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Enamel organ;
RA   Ding Y., Xue J., Bartlett J.D.;
RT   "T-plasminogen activator in tooth tissues.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR   EMBL; AF364605; AAM00297.1; -.'
DR   HSSP; P00761; 1AN1.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctn1.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   ProDom; PD003395; Kringle; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.

```
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease
SQ   SEQUENCE   562 AA;  63668 MW;  F9E6B4C77CB101E8 CRC64;

Query Match             81.0%;  Score 1669;  DB 6;  Length 562;
Best Local Similarity   82.0%;  Pred. No. 8.9e-158;
Matches  292;  Conservative  27;  Mismatches  37;  Indels   0;  Gaps   0;

Qy   21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHN 80
            :  :|| ||||||||||||||:| |:||:|| ||||||||:|| |||||||||:|| |||||
Db  207 ACTKEKEBCYTGKGLDYRGTRSLTMSGAFCLPWNSLVLMGKIYTAWNSNAQTLGLGKHNY 266

Qy   81 CRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQA 140
            ||||||| |||||||||||||  ||| ||||:|||||: |: :||||||||:|.|:||
Db  267 CRNPDGDTQPWCHVLKDHKLTWEYCDLLPQCVTCGLRQYKEPQFRIKGGLYADITSHPWQA 326

Qy  141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQK 200
            ||||||||||||||||||||:||||||||||||||||||||  ||:|||||||||||: |
Db  327 AIFVKNRRSPGERFLCGGVLISSCWVLSAAHCFQERFPPHHVRVVLGRTYRLVPGEEEQA 386

Qy  201 FEVEKYIVHKEFDDDTYNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECEL 260
            |||||||||||||||||||||||||||||||:||||||:| |||.|||:|||||||||
Db  387 FEVEKYIVHKEFDDDTYNDIALLQLKSDSLTCAQESDAVRTVCLPEANLQLPDWTECEL 446

Qy  261 SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320
            |||||||: :|||||||||||||||||||||||||:|.|:||:|||||||||||  |||
Db  447 SGYGKHEASSPFFSERLKEAHVRLYPSSRCTSKHLFNKTITNNMLCAGDTRSGGDNANLH 506

Qy  321 DACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376
            |||||||||||||  ||:||||||||||||||||||||||||||||||||||:|
Db  507 DACQGDSGGPLVCMKGNHMTLVGVISWGLGCGQKDVPGVYTKVTNYLNWIRDNTRP 562

RESULT 4
Q91VP2
ID   Q91VP2       PRELIMINARY;      PRT;   559 AA.
AC   Q91VP2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to plasminogen activator, tissue.
GN   PLAT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR   EMBL; BC011256; AAH11256.1; -.
DR   HSSP; P00761; 1AN1.
DR   MGD; MGI:97610; Plat.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctnl.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
```

```
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease.
SQ   SEQUENCE   559 AA;  63122 MW;  8CCE82BDB94514D9 CRC64;

Query Match             77.7%;  Score 1601;  DB 11;  Length 559;
Best Local Similarity   79.5%;  Pred. No. 5.5e-151;
Matches  283;  Conservative  34;  Mismatches  37;  Indels   2;  Gaps   2;

Qy   21 AASEGNS-DCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHN 79
            || :|||  |  | |||||||:||| || :|:|||||  |:|| |||| :|:||||| ||
Db  204 ACPKGKSEDCYVGKGVTYRGTHSLTTSQASCLPWNSIVLMGKSYTAWRTNSQALGLGRHN 263

Qy   80 YCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQPRIKGGLFADIASHPWQ 139
            ||||||||| |||||||:|  |||||||  ||:|||||: |: :|||||||  |:|||||
Db  264 YCRNPDGDARPWCHVMKDRKLTWEYCDMSPCSTCGLRQYKRPQFRIKGGLYTDITSHPWQ 323

Qy  140 AAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQ 199
            | |:| |:|:|||||||||||||||||:|||||||:|:||||  :|:|||||||||||||
Db  324 APIFVKNRSPGERFLCGGVLISSCWVLSAAHCFLERPPNHLKVVLGRTYRVVPGEEEQ 383

Qy  200 KFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECE 259
            ||||||||||:|||||||||||||||||:|:|||||||  |:|||||| |||:|||||||
Db  384 TFEIEKYIVHEEFDDDTYDNDIALLQLRSQSKQCAQESSVGTACLPDPNLQLPDWTECE 443

Qy  260 LSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANL 319
            ||||||||| |||:|||||||||||||||||||||| :| |||||||||||||||| :|
Db  444 LSGYGKHEASSPFFSDRLKEAHVRLYPSSRCTSQHLFNKTVTNNMLCAGDTRSGGNQ-DL 502

Qy  320 HDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMR 375
            |||||||||||||  :  : ||| ||||||||||||||||||||||:|:||||:||
Db  503 HDACQGDSGGPLVCMINKQMTLTGIISWGLGCGQKDVPGVYTKVTNYLDWIHDNMK 558

RESULT 5
Q8MKB1
ID   Q8MKB1       PRELIMINARY;      PRT;   564 AA.
AC   Q8MKB1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Tissue-type plasminogen activator.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR   EMBL; AY029518; AAK40240.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctnl.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 1.
```

```
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM001181; EGF; 1.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease.
SQ   SEQUENCE   564 AA;  62726 MW;  459DBBAC6D4A937C CRC64;

Query Match             74.5%;
Best Local Similarity   75.7%;   Score 1536;  DB 6;  Length 564;
Matches  274;   Conservative  29;   Mismatches  59;   Indels  0;   Gaps  0;

Qy    15  FATVAQAASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALG  74
Db   203  FCSTPACSKEKNGNCYLGKGQAYRGTHSLTTSGASCLPWNSMLLVGEKYTARQSNAEALG 262

Qy    75  LGKHNYCRNPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIA 134
Db   263  LGKHNYCRNPDGDSKPWCHVLKNRKLVIEYCDVPQCATCGLRQDKQPQFRIKGGLFTDIT 322

Qy   135  SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVP 194
Db   323  AHPWQAAIFTNNRRSPGERFLCGGILINSCWVLSAAHCFLERFPQQKLRVIILGRTYPLVS 382

Qy   195  GEEEQKFEVEKYIVHKEFDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPD 254
Db   383  AEEEQIFEVEQPILHERFDEGTYDNDIALLKLKSTSGGSCAQESQAVRLVCLPDASLQLPD 442

Qy   255  WTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGG 314
Db   443  WTECELSGYGKHEEFSPVFSEQLKEAHVRLYPSSRCTPQQLKNRTVTGNMLCAGDTRSGG 502

Qy   315  PQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNM 374
Db   503  AQVNLHDACQGDSGGPLVCMTDGHMTLIGIISWGLGCGQKDVPGVYTKVVNYLGWIQQHV 562

Qy   375  RP 376
Db   563  GP 564
```

```
RESULT 6
Q8MIL0
ID   Q8MIL0          PRELIMINARY;     PRT;   433 AA.
AC   Q8MIL0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Urokinase-type plasminogen activator.
GN   PLAU.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22155945; PubMed=12149463;
RA   Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA   Dichek D.A.;
```

```
RT   "Increased expression of urokinase during atherosclerotic lesion
RT   development causes arterial constriction and lumen loss, and
RT   accelerates lesion growth";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; AY122285; AAM83187.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ   SEQUENCE   433 AA;  48375 MW;  65E64F3641554980 CRC64;

Query Match             37.3%;     Score 769;  DB 6;  Length 433;
Best Local Similarity   43.1%;     Pred. No. 4.9e-68;
Matches  157;   Conservative  52;   Mismatches  127;   Indels  28;   Gaps  8;

Qy    29  CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA  88
Db    72  CYHGDGHSYRGKANTDIMDRPCLAWNSANVLTKTYHAHRPDALQLGLGKHNYCRNPDHQR 131

Qy    89  KPWCHVLKNRRLTWEYCDVPSCST------------CGLRQYSQPQFRIKGGLFADIA   134
Db   132  RPWCYVQVGLKQLIQECKVHDCSSGKKPALPPGKLEFQCG-QKALRPRKIIGGEFTTIE  190

Qy   135  SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRT--YRV 192
Db   191  NQPWFAAIYRRH-RGGSVTYVCGGSLISPCWVSATHCFINHQKKEDYIVYLGRSRLNSM 249

Qy   193  VPGEEEQKFEVEKYIVHKEFDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADL 250
Db   250  TPG--EMKFEVEQLILHBGYSADTLAHHNDIALLKILSNNGQCAQPSRSIQTICLPPWNA 307

Qy   251  QLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDT 310
Db   308  DPNFGTSCEITGFGKENSTDYLYPEQLKMTVKLVSYQECQQPHYYGSEVTKMLCAAD-   366

Qy   311  RSSGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWI 370
Db   367  ----PQWET-DSCQGDSGGPLVCSVQGRMLLTGIVSWGRGCALKNKPGVYTRVSRPLPWI 425

Qy   371  RDNM 374
Db   422  RSHI 425
```

```
RESULT 7
Q8MHY7
ID   Q8MHY7          PRELIMINARY;     PRT;   433 AA.
AC   Q8MHY7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Urokinase-type plasminogen activator.
GN   UROKINASE.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
```

RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yano W., Watanabe M.;
RT   "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT   complete cds.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; AY029517; AAK40239.1; -.
DR   EMBL; AB087224; BAC02685.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ   SEQUENCE   433 AA;   48444 MW;   6DD35A37101OA6EE CRC64;

Query Match   37.2%; Score 767; DB 6; Length 433;
Best Local Similarity   43.1%; Pred. No. 7.8e-68;
Matches 157; Conservative   52; Mismatches 127; Indels 28; Gaps 8;

Qy   29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
         ||::|:|||| : ||||||||
Db   72 CTHGDGHSYRGKANTDIMDRPCLAWNSANVLTKTYHAHRPDALQLGLGKHNYCRNPDHQR 131

Qy   89 KPWCHVLKNRRLTWEYCDVPSCST---------CGLRQYSQPQFRIKGGLFADIA 134
         ||:|::||| || : |:|:|||||||||||
Db   132 RPWCYVQVGLKQLIQECKVHDCSSGKKPALPPGKLEFQCG--QKALRPRFKIIGGEFTIIE 190

Qy   135 SHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRT--YRV 192
         :||:||| ||||||||||||||||| :
Db   191 NQPWFAAIYRRH-RGGSVTYVCGGSLISPCWVVSATHCFINHQKKEDYIVYLGRSRLNSM 249

Qy   193 VPGEEEQKFEVEKYIVHKEFDDT--YDNDIALLQLKSDSSRCAQESSVVRTVCLPPADL 250
         |:||| |:|:   :|||||||||:   ::||| :::||
Db   250 TPG--EMKFEVEQLIIHEGYRADTLAHHNDIALLKILSNNGQCAQPSRSIQTICLPPWNA 307

Qy   251 QLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDMLCAGDT 310
         ||::|:| ||| |:|| :|| |: ||||:|||||||||||
Db   308 DPNFGTSCEITGFGKENSTDYLYPEQLKMTVVKLVSYQBCQQPHYGSEVTTKMLCAAD- 366

Qy   311 RSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCCGQKDVPGVYTKVTNYLDWI 370
         ||||| |:|||||||| |:|||||||:|||:|||:||||| :|:||:|| :|||
Db   367 ----PQWET-DSCQGDSGGPLVCSVQGRMTLTGIVSWGRGCALKNKPGVYTRVSRFLPWI 421

Qy   371 RDNM 374
         ::|:
Db   422 RSHI 425

RESULT 8
Q90675
ID   Q90675   PRELIMINARY;   PRT; 202 AA.
AC   Q90675;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update).
DE   Tissue-type plasminogen activator (fragment).
GN   TPA.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=White Leghorn;
RX   MEDLINE=97199025; PubMed=9047000;
RA   Johnson A.L., Bridgham J.T., Anthony R.V.;
RT   "Expression of avian urokinase and tissue-type plasminogen activator
RT   messenger ribonucleic acid during follicle development and atresia.";
RL   Biol. Reprod. 56:581-588(1997).
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR   EMBL; U31988; AAA74955.1; -.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00130; KR; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT   NON_TER   1   1
FT   NON_TER   202   202
SQ   SEQUENCE   202 AA;   23148 MW;   049DB42941D83AFC CRC64;

Query Match   37.0%; Score 763; DB 13; Length 202;
Best Local Similarity   77.4%; Pred. No. 7.1e-68;
Matches 130; Conservative   16; Mismatches 22; Indels 0; Gaps 0;

Qy   23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCR 82
         :||:::|||||||||||||||||||||||||||||||||:||||||||||||||||||:||
Db   35 AEEDGDCYTGNGLAYRGTRSRTKSGFSCLPWNPVFLTSKIYTALEEQRRALGLGKHNHCR 94

Qy   83 NPDGDAKPWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAI 142
         |||||||||||||||||||||||||||||| :|:|||||||||||||||||:||||||||
Db   95 NPDGDAQPWCHVWKDRQLTWEYCDVPVPQCVTCGLRQYKRPQFRIKGGLFADITSHPWQAAI 154

Qy   143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHHLTVILGRTY 190
         |:|:|:|||||||||||||||||||||||||||||:||||||||||||
Db   155 FVKNRRAPGQRFLCGGILISSCWVLSAAHCFQERYPPHHLTVILGRTY 202

RESULT 9
Q8VCS4
ID   Q8VCS4   PRELIMINARY;   PRT; 653 AA.
AC   Q8VCS4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 70.6 kDa protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Muridae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; BC019376; AAH19376.1; -.

```
DR   HSSP; P00761; 1AN1.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctnl.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00013; FNTYPEII.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPC; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS00070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW   Kringle; Protease; Serine protease.
SQ   SEQUENCE   653 AA;  70553 MW;  FE18D90174ED6FDD CRC64;

Query Match             36.0%;  Score 742.5;  DB 11;  Length 653;
Best Local Similarity   40.1%;  Pred. No. 3.8e-65;
  Matches  153;  Conservative  58;  Mismatches  118;  Indels  53;  Gaps   9;

Qy         29 CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDA 88
              :  ||||| :: |||||||||||||||||||||||||||||||||||| ::|||
Db        283 CFLGNGTEYRGVASTAASGLSCLAWNSDLLYQELHVDSVAAAVLLGLGPHAYCRNPDKDB 342

Qy         89 KPWCHVLKNRRLTWEYCDVPSCS--------------------------TCGLRQ---YS 119
              :||:||| :||||||||| : ||                          ||||||   ::
Db        343 RPWCYVVKDNALSWEYCRLTACESLARVHSQSPEILAALPESAPAVRPTCGKRHKKRTFL 402

Qy        120 QPQPRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPP 179
              ::   || : |||||||||||||| :||:     ::::||||||||  :  :|||||:|:
Db        403 RP--RIIGGSSLPGSHPWLAAIYI------------GNSF-CAGSLVHTCWVVSAAHCFANSPPR 453

Qy        180 HHLTVILGRTYRVVPGEEEQKFEVEKYI---VHKEFDDDTYNDIALLQLKSDSSRCAQE 236
              :|||||||||||||||||||||||||||    ||::||  :||:|||||||||:||||||
Db        454 DSITVVLGQHFFNRTTDVTQTFGIEKVVPYTLYSVFNPN.-NHDLVLRLKKKGERCAVR 511

Qy        237 SSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLL 296
              :|| |:||| |: ||| | : ||||::|||:| |||  :||:::|::||:||||:|||||
Db        512 SQFVQPICLPEAGSSFPGHKCQIAGWGHMDENVSSYSNSLLEALVPLVADHKCSSPEVY 571

Qy        297 NRTVTDNMLCAG--DTRSGGPQANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCGQK 354
              :: ||||||| :  :  ||||: :|||||||||||||:| : ||||||||||||||:|||
Db        572 GADISPNMLCAGYFDCKS------DACQGDSGGPLVCEKNGVAYLYGIISWGDGCGRL 623

Qy        355 DVPGVYTKVTNYLDWIRDNMRP 376
              :|||||| ||||| ||||:|||:
Db        624 NKPGVYTRVANYDWINDRIRP 645


RESULT 10
ID   Q14520;        PRELIMINARY;        PRT;   560 AA.
AC   Q14520; O00663;
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   HGF activator like protein (Hyaluronan binding protein 2).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kitamura N.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96425001; PubMed=8827452;
RA   Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA   Tomita M.;
RT   "Purification and characterization of a novel hyaluronan-binding
RT   protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT   serine protease domain, similar to hepatocyte growth factor
RT   activator.";
RL   J. Biochem. 119:1157-1165(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon, and Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; D49742; BAA08576.1; -.
DR   EMBL; S83182; AAB46909.1; -.
DR   EMBL; BC031412; AAH31412.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.033; -.
DR   Genew; HGNC:4798; HABP2.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 3.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; EGF; 1; 3.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS00070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease
SQ   SEQUENCE   560 AA;  62671 MW;  5C19072307B4ACD4 CRC64;

Query Match             34.4%;  Score 709.5;  DB 4;  Length 560;
Best Local Similarity   39.1%;  Pred. No. 6.1e-62;
  Matches  148;  Conservative  62;  Mismatches  128;  Indels  41;  Gaps  10;

Qy         25 GNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNP 84
              |: |||  |||||||||||||||:|:|||||||||||||||| :: |:||||
Db        190 GSDDCYVGDGYSYRGKMNRTVNQHACLYWNSHLLQENYNMFMEDAETHGIGEHNFCRNP 249

Qy         85 DGDAKPWCHV-LKNRRLTWEYCDVPSCS-----------------TCGLRQYSQP 121
              :|||||:|:  ::|||||:|||||:|||                 :|:|:|:||:
Db        250 DADEKPWCFIKVTNDKVKWEYCDVSACSAQDVAYPEESPTEPSTKLPGFDSCGKTEIAER 309

Qy        122 QF-RIKGGLFADIASHPWQAAIFAK--HRRSPGERFLCGGILISSCWILSAAHCFQERFP 178
                :  |||: :|||||||||||||    ||| ||| :||||:|||| :  :|||||:||
Db        310 KIKRIYGGFKSTAGKHPWQASLQSSLPLTISMPQGHFCGGALIHPCWVLTAAHCTD--IK 367

Qy        179 PHHLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFD--DDTYDNDIALLQLKSDSSRCAQE 236
```

```
Db    368  TRHLKVVLGDQDLKKEEFHEQSFRVEKIFKYSHYNERDEIPHNDIALLKLKPVDGHCALE  427
             ||  ||:||       ||||  ||  : ::  |||||  ||||:||:|||:||
Db    207  DGDHKPWCFVKVNSEKVKWEYCDVTVCVPVPDTNPVESLLEPVMELPGPESCGKTEVAEH  266

Qy    237  SSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLL  296
             : |:|||      ||||||  :||| ||  ||| | :    ||||||||||||:
Qy    122  QF-RIKGGLFADIASHPWQAAIFAK--HRRSPGERFLCGGILISSCWIISAAHCFQERFP  178

Db    428  SKYVKTVCLP--DGSFPSGSECHISGWGVTETGKG--SRQLLDAKVKLIANPLCNSRQLY  483
             : |||:|||    |:| ||  :||  ||:||:||    ||| ||  ||| ||:||||:
Db    267  AVKRIYGGFKSTAGKHFWQVSLQTSLPLTTSMPQGHFCGGALLHPCWVLTAAHCTD--IN  324

Qy    297  NRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDV  356
             ||| ||| |||||||||||||||||| |||||||||||   |||||||||||| ||:
Qy    179  PHHLTVILGRTYRVVPGEEEQKPVEKYIVHKFPD--DDTYDNDIALLQLKSDSSRCAQE  236

Db    484  DHMIDDSMICAGNLQKPG-----QDTCQGDSGGPLTCEKDGTYYVVGIVSWGLECGKR--  536
             ::|:||| ||| ||||     ||||||||||||| |:||||| :  ||| |||| |
Db    325  TKHLKVVLGDQDLKKTESHEQTFRVEKLKYSQYNERDEIPHNDIALLKLKPVGGHCALE  384

Qy    357  PGVYTKVTNYLDWIRDNMR  375
             |||||||| ||||||:|
Qy    237  SSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLL  296

Db    537  PGVYTQVTKFLNWIKATIK  555
                                Db    385  SRYVKTVCL-PSD-PPPSGTECHISGWGVTETGEG--SRQLLDAKVKLIANPLCNSRQLY  440

Qy    297  NRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDV  356
                                                 ----SDTCQGDSGGPLTCEKDGTYYVVGIVSWGQECGKK--  493
                                            Db    441  DHTIDDSMICAGNLQKPG

                                            Qy    357  PGVYTKVTNYLDWIRDNMR  374
                                                       |||||||| ||||||:|
                                            Db    494  PGVYTQVTKFLNWIKTTM  511
```

```
RESULT 11
Q8KOD2
ID   Q8KOD2         PRELIMINARY;        PRT;     517 AA.
AC   Q8KOD2;
DT   01-OCT-2002 (TREMBLrel. 22, Created)
DT   01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; BC031775; AAH31775.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW   Kringle; Protease; Serine protease.
SQ   SEQUENCE   517 AA;  57326 MW;  3B5SA42O35A5EA59 CRC64;

Query Match             33.9%;  Score 698.5;  DB 11;  Length 517;
Best Local Similarity   39.7%;  Pred. No. 6.9e-61;
Matches  150;  Conservative  54;  Mismatches 133;  Indels  41;  Gaps  11;
```

```
Qy    25   GNSDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNP  84
             ||||    || ||| |||| ||||||||||||  |:||  :::|: |||| :||||
Db    147  GPDDCYYGDGYSYRGKVSKTVNQNPCLYWNSHLLQETYNMFMEDAETHGIAEHNFCRNP  206

Qy    85   DGDAKPWCHV-LKNRLTWEYCDVPSC------------------STCGLRQYSQP  121
             ||||||| |  ||||||||||||| |                  :||   :: ::
```

```
RESULT 12
Q8IZZ5
ID   Q8IZZ5         PRELIMINARY;        PRT;     615 AA.
AC   Q8IZZ5;
DT   01-MAR-2003 (TREMBLrel. 23, Created)
DT   01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE   Coagulation factor XII-Mie.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT   "Molecular chracterization of coagulation factor XII-Mie.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB095845; BAC23095.1; -.
SQ   SEQUENCE   615 AA;  67735 MW;  030508870A0C7EDB CRC64;

Query Match             32.4%;  Score 667.5;  DB 4;  Length 615;
Best Local Similarity   36.2%;  Pred. No. 1.1e-57;
Matches  150;  Conservative  48;  Mismatches 133;  Indels  83;  Gaps  7;
```

```
Qy    27   SDCYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDG  86
             ||  || |||| |||| |||||||||||||| ||||   :::|:    ||:|:|||||
Db    215  ASCYDGRCLSYRGLARTTLSGAPCQPWASEATYRNVTABQ---ARNWGLGGHAFCRNPDN  271

Qy    87   DAKPWCHVLKNRRLTWEYCDVPSCST---------  112
             || |||  ||||||||||||| |               |
Db    272  DIRPWCFVLNRDRLSWEYCDLAQCQTPTQAAPTPVSPRLHVPLMPAQPAPPKPQPTTRT  331

Qy    113  ------------------------CG--LRQYSQPQFRIKGGLFADIASHPWQAIF  143
                                     ||   ||  || |  ||||||| :|  || ||  :
Db    332  PPQSQTPGALPAKREQPPSLTRNGPLSCGQRLRKSLSSMTRVVGGLVALRGAHPYIAALY  391

Qy    144  AKHRRSPGERFLCGGILISSCWIISAAHCFQERPPPHHLTVILGRTYRVVPGEEEQKPEV  203
             |||| ||||||||||||  ||||||||||||   || |||:|||||||| ||||||||
Db    392  WGHS------------FCAGSLIAPCWVLTAAHCLQDRPAPEDLTVVLGQERRNHSCEPCQTLAV  444

Qy    204  EKYIVHKFDDTYDNDIALLQLKSDS-SRCAQESSVVRTVCLPPADLQLPDWTECELSG  262
             || ||||||||||||||||| ||||   |  || | ||| ||||:||||||||||:
Db    445  RSYRLHEAFSPVSYQHDLALLRLQEDADGSCALLSPVVQPVCLPSGAARPSETTLCQVAG  504

Qy    263  YGHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA  322
             || | || |:||   |||  |   || | |:   :||  | |||  |||| ----TDA
Db    505  CGHQFEGAEEYASFLQEAQVPFLSLERCSAPDVHGSSILPGMLCAGFLEGG-----TDA  558

Qy    323  CQGDSGGPLVCLN---DGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDN  373
             ||||||||| |     ||||||||||||||||||||||  ||||||| |||||:
```

```
Qy    121  STCGLRQYSQP
Db    375  PGVYTKVTNYLDWIRDNMR
Db    555  PGVYTQVTKFLNWIKATIK
```

Db   559 CQGDSGGPLVCEDQAAERRLTLQGIISWGSGCGDRNKPGVYTDVAYYLAWIREH 612

RESULT 13
035727
ID   035727;   PRELIMINARY;   PRT;   597 AA.
AC   035727;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Factor XII.
GN   F12.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Schloesser M., Schwager S., Engel W.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; X99571; CAA67891.1; -.
DR   HSSP; P00760; 1AQ7.
DR   MEROPS; S01.211; -.
DR   MGD; MGI:1891012; F12.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctnl.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00013; FNTYPEII.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   ProDom; PD000395; Kringle_1; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease;
SQ   SEQUENCE   597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match   30.8%; Score 634.5; DB 11; Length 597;
Best Local Similarity   36.4%; Pred. No. 2.1e-54;
Matches 144; Conservative 48; Mismatches 137; Indels 67; Gaps   9;

Qy   29 CYFGNGSAYRGTHSLTESGASCLPWNSMLIGKVY-TAQNPSAQALGLGKHNYCRNPDGD 87
           ||:|||:||| : ||||||||:|| | :||| :: ||| |||||||||
Db   217 CYEGRGLSYRGQAGTTQSGAPCQRWT----VEATYRNMTEKQALSWGLGHHAFCRNPDND 272

Qy   88 AKPWCHVLKNRRLTWEYCDVPSCST------------------ 112
           ||:|||| |||||| | : ::
Db   273 TRPWCFVWSGDRLSWDYCGLEQCQTPTAPLVVPESQESPSQAPSLSHAPNDSTDHQTS 332

Qy   113 ------------CG--LRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILIS 162
                       || ||:|| ||:||||:|| |||| || | ||: ::
Db   333 LSKTNTMCGQGRFRKGLSSFMRVVGGLVALPGSHPYIAALYW-----GNNF-CAGSLIA 385

Qy   163 SCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIA 222
           ||||| :|| ::| |||| | || |||:|| | | ::||:|:||:|
Db   386 PCWVLTAAHCLQNRPAPEELTVVLGQDRHNQSCEWCQTLAVRSYRLHEGFSSITYQHDLA 445

Qy   223 LIQL-KSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAH 281
           ||:| :| :|| || | |:|||: | | || | |:| | || | |:|||
Db   446 LLRLQESKTNSCAILSPHVQPVCLPSGAAPPSETVLCEVAGWGHQLEGAEEYSTFLQEAQ 505

Qy   282 VRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVC---LNDGR 338
           ||| | :|| |||| || |||| ||||||| :||||||||||||| :
Db   506 VPFIALDRCSNSNVHGDAILPGMLCAGFLEGG------TDACQDSGGPLVCEEGTAEHQ 559

Qy   339 MTLVGIISWGLGCGQKDVPGVVTKVTNYLDWIRDNM 374
           :|| || |:|| || | |||:||||| |:: :;
Db   560 LTLRGVISWGSGCGDRNKPGVVTDVANYLAWIQKHI 595

RESULT 14
097507
ID   097507;   PRELIMINARY;   PRT;   616 AA.
AC   097507;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   FXII.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Takahashi T., Kihara T.;
RT   "Porcine liver factor XII.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR   EMBL; AB022426; BAA37148.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.211;-.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000083; Fibrnctnl.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00039; fn1; 1.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00051; kringle; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00013; FNTYPEII.
DR   PRINTS; PR00018; KRINGLE.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   ProDom; PD000395; Kringle; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00058; FN1; 1.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00130; KR; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.

```
DR   PROSITE; PS00070; KRINGLE_2; 1.
DR   PROSITE; PS00240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW   Serine protease.
SQ   SEQUENCE   616 AA;  68012 MW;  4C5FE3D71EBBD1A9 CRC64;
```

```
Query Match                30.6%;   Score 630;  DB 6;  Length 616;
Best Local Similarity     35.8%;   Pred. No. 6e-54;
Matches 149;  Conservative  48;  Mismatches 129;  Indels  90;  Gaps  11;
```

```
Qy    29  CYFGNGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSA-QAL--GLGKHNYCRNPD  85
              |:| |||||||||||||||||||||||| :|||   |||  ::|||||
Db   217  CYSDRGLSYRGMAQTTLSGAPCQPWASEA------TYWNMTAQALNWGLGDHAFCRNPD  270

Qy    86  GDAKPWCHVLKNRRLTWEYCDVPSCST------------------------  112
              |:|| ||||||||  :|:||:|
Db   271  NDTRPWCFVWRGDQLSWQYCRLARCQAPIGEAPPILTPTQSPSEHQDSPLLSREPQPTTQ  330

Qy   113  ------------------CG--LRQYSQPFRIKGGLFADIASHPWQAAIF  143
                                || ||||  ||||: ||::|
Db   331  TPSQNLTSAWCAPPEQRGPLPSAGLVGCGQRLRKRLSSLNRIVGGLVALPGAHPYIAALY  390

Qy   144  AKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFEV  203
              |: |||  |: ||||: ||  |:|:|
Db   391  W------GQNF-CAGSLIAPCWVLTAAHCLQNRPAPEELTVVLGQDRHNQSCEQCQTLAV  443

Qy   204  EKYIVHKEFDDDTYDNDIALLQLKSDSSR-CAQESSVVRTVCLP--PADLQLPDWTECEL  260
              |||: ||||||||||: |||| :||||   |: || :|| :|  ||||||||||||
Db   444  RSYRLHESYSPKTYQHDLALVRLKETADGCCAHPSPFVQPVCLPRSVASSARPEGALCEV  503

Qy   261  SGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH  320
              :|: ||   ||:|  :|:||   |:|| ||
Db   504  AGWGHQFEGAEBYSSFLQEAQVPLISPERCSAADVHGAAFTPGMLCAGFLEGG------T  557

Qy   321  DACQGDSGGPLVCLN--DGRMTLVGIISWGLGCGQKDVPGVVTKVINYLDWIRDN  373
             |||||||||||||    :|:||||||||||||  ||: |||||||||||| :||
Db   558  DACQGDSGGPLVCEDETAERQLVLRGIVSWGSGCGDRLKPGVVTDVANYLAWIQEH  613
```

```
Query Match                29.5%;   Score 607;  DB 11;  Length 812;
Best Local Similarity     36.7%;   Pred. No. 1.7e-51;
Matches 139;  Conservative  54;  Mismatches 132;  Indels  54;  Gaps  13;
```

```
Qy    17  TVAQ--AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQAL  73
              |:|   ||||||||||||||||||||||:: :|:
Db   466  TVSQEPSGPSDSETDCMYGNGKDYRGKTAVTAAGTPCQGW--------AAQEPHRHSI  515

Qy    74  --------GLGKHNYCRNPDGDAK-PWCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQF  123
                      :|||:|||||| || ||:||||| :||||||  ||:|| |:  :|:|
Db   516  FTPQTNPRAGLEK-NYCRNPDGDVNGPWCYTTNPRKL-YDYCDIPLCASASSFECGKPQV  573

Qy   124  ------RIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQER  176
                    |:||| ||| ||||||| | |  : ||||||||||| |||||||||||
Db   574  EPKKCPGRVVGGCVANPHSWPWQISL---RTRFTGQHF-CGGTLIAPBWVLTAAHCLEKS  629

Qy   177  FPPHHLTVILGRTYRVVPGEEEQKFEVEKYIVHKEFDDDTYDNDIALLQLKSDSSRCAQE  236
              |: | ||||||| ||  :| |::|
Db   630  SRPEFYKVILGAHBEYIRGSDVQEISVAKLIL------EPNNRDIALLKL----SRPATI  679

Qy   237  SSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLL  296
              :: ||  ||   | ||:|| :|| :||| : || | |   :
Db   680  TDKVIPACLPSPNVMVADRTICYITGWG--ETQGTFGAGRLKEAQLPVIENKVCNRVEYL  737

Qy   297  NRTVTDNMLCAGDTRSGGPQANLHDACQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDV  356
              |: |||||||| :||||||||||:||||||||||||:  |:| ||||||||||
Db   738  NNRVKSTELCAGQLAGG------VDSCCQDSGGPLVCFEKDKYILQGVTSWGLGCARPNK  791

Qy   357  PGVYTKVINYLDWIRDNMR  375
              |||| |:||| ||:|
Db   792  PGVYVRVSRFVDWIEREMR  810
```

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 19:53:48 ; Search time 277.124 Seconds
                                            (without alignments)
                                            9743.047 Million cell updates/sec

Title: US-09-987-455-3
Perfect score: 66
Sequence: 1 atgaaaagacagctatcgc........ctaccgtggcccaggcggcc 66

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : GenEmbl:*
         1: gb_ba:*
         2: gb_htg:*
         3: gb_in:*
         4: gb_om:*
         5: gb_ov:*
         6: gb_pat:*
         7: gb_ph:*
         8: gb_pl:*
         9: gb_pr:*
        10: gb_ro:*
        11: gb_sts:*
        12: gb_sy:*
        13: gb_un:*
        14: gb_vi:*
        15: em_ba:*
        16: em_fun:*
        17: em_hum:*
        18: em_in:*
        19: em_mu:*
        20: em_om:*
        21: em_or:*
        22: em_ov:*
        23: em_pat:*
        24: em_ph:*
        25: em_pl:*
        26: em_ro:*
        27: em_sts:*
        28: em_un:*
        29: em_vi:*
        30: em_htg_hum:*
        31: em_htg_inv:*
        32: em_htg_other:*
        33: em_htg_mus:*
        34: em_htg_pln:*
        35: em_htg_rod:*
        36: em_htg_mam:*
        37: em_htg_vrt:*
        38: em_sy:*
        39: em_htgo_hum:*
        40: em_htgo_mus:*
        41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	66	6	AX430829	AX430829 Sequence
2	66	100.0	66	6	AX431269	AX431269 Sequence
3	66	100.0	66	6	AX431272	AX431272 Sequence
4	66	100.0	705	12	HSIGKLC28	X95748 Synthetic c
5	66	100.0	705	12	HSIGKLC5	X95747 Synthetic c
6	66	100.0	711	12	HSIGKLC14	X95749 Synthetic c
7	66	100.0	711	12	HSIGKLG31	X95750 Synthetic c
8	66	100.0	1128	6	AX431268	AX431268 Sequence
9	66	100.0	1128	6	AX431271	AX431271 Sequence
10	66	100.0	3394	12	AF268280	AF268280 Phagemid
11	66	100.0	3450	6	AR241678	AR241678 Sequence
12	66	100.0	3758	12	AF268281	AF268281 Phagemid
13	66	100.0	5149	6	AX591634	AX591634 Sequence
14	66	100.0	5149	6	AX600065	AX600065 Sequence
15	66	100.0	5683	6	AX554424	AX554424 Sequence
16	66	100.0	6122	6	AX554413	AX554413 Sequence
17	66	100.0	6122	6	AX554422	AX554422 Sequence
18	66	100.0	6122	6	AX642149	AX642149 Sequence
19	66	100.0	6125	6	AX554420	AX554420 Sequence
20	64.4	97.6	5885	1	STAFRA	L42764 Staphylococc
21	61.2	92.7	580	6	AX742885	AX742885 Sequence
22	61.2	92.7	580	6	AX743509	AX743509 Sequence
23	60.2	91.2	74	6	E02814	E02814 DNA sequenc
24	60.2	91.2	74	6	E03955	E03955 DNA encodin
25	60.2	91.2	103	6	AX236721	AX236721 Sequence
26	60.2	91.2	249	6	I08200	I08200 Sequence 1
27	60.2	91.2	249	6	I52035	I52035 Sequence 1
28	60.2	91.2	438	12	SYNOMPARNA	M62710 Synthetic g
29	60.2	91.2	548	6	AR179433	AR179433 Sequence
30	60.2	91.2	548	6	AR282580	AR282580 Sequence
31	60.2	91.2	548	6	AX022441	AX022441 Sequence
32	60.2	91.2	548	6	AX030798	AX030798 Sequence
33	60.2	91.2	548	6	AX034619	AX034619 Sequence
34	60.2	91.2	652	6	E01685	E01685 hybrid DNA
35	60.2	91.2	739	1	AF234271	AF234271 Shigella
36	60.2	91.2	746	1	AF234270	AF234270 Salmonell
37	60.2	91.2	1400	1	STOMPA	X02006 Salmonella
38	60.2	91.2	1697	1	AF234269	AF234269 Escherich
39	60.2	91.2	2270	1	AX191724	AX191724 Sequence
40	60.2	91.2	2271	1	ECOMPA	V00307 E. coli omp
41	60.2	91.2	6477	6	A08458	A08458 pinf 4-49 m
42	60.2	91.2	7499	12	ASAK3121	AJ223121 Artificia
43	60.2	91.2	7500	12	ASAK3122	AJ223122 Artificia
c 44	60.2	91.2	11639	1	AE000198	AE000198 Escherich
c 45	60.2	91.2	11641	1	AE005286	AE005286 Escherich

ALIGNMENTS

RESULT 1
AX430829
LOCUS   AX430829   66 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION   Sequence 1 from Patent WO0240696.
ACCESSION   AX430829
VERSION   AX430829.1 GI:21655908
KEYWORDS   .
SOURCE   Escherichia coli
ORGANISM   Escherichia coli
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Escherichia.
REFERENCE   1
AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
TITLE   Methods for large scale protein production in prokaryotes
JOURNAL   Patent: WO 0240696-A 1 23-MAY-2002;

Pred. No. is the number of results predicted by chance to have a

```
FEATURES      BOEHRINGER INGELHEIM INT (DE)
     source            1..66
                       /organism="Escherichia coli"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:562"
BASE COUNT      14 a      18 c      21 g      13 t
ORIGIN

Query Match            100.0%;  Score 66; DB 6; Length 66;
Best Local Similarity  100.0%;  Pred. No. 1.4e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66


RESULT 2
AX431269
LOCUS         AX431269              66 bp    DNA       linear   PAT 28-JUN-2002
DEFINITION    Sequence 3 from Patent WO0240650.
ACCESSION     AX431269
VERSION       AX431269.1  GI:21656151
KEYWORDS      .
SOURCE        Escherichia coli
  ORGANISM    Escherichia coli
              Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
              Enterobacteriaceae; Escherichia.
REFERENCE     1
  AUTHORS     Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE       Methods for large scale production of recombinant dna-derived tpa
              or k2s molecules
  JOURNAL     Patent: WO 0240650-A 3 23-MAY-2002;
              BOEHRINGER INGELHEIM INT (DE)
FEATURES      Location/Qualifiers
     source            1..66
                       /organism="Escherichia coli"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:562"
BASE COUNT      14 a      18 c      21 g      13 t
ORIGIN

Query Match            100.0%;  Score 66; DB 6; Length 66;
Best Local Similarity  100.0%;  Pred. No. 1.4e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66


RESULT 3
AX431272
LOCUS         AX431272              66 bp    DNA       linear   PAT 28-JUN-2002
DEFINITION    Sequence 6 from Patent WO0240650.
ACCESSION     AX431272
VERSION       AX431272.1  GI:21656154
KEYWORDS      .
SOURCE        Escherichia coli
  ORGANISM    Escherichia coli
              Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
              Enterobacteriaceae; Escherichia.
REFERENCE     1
  AUTHORS     Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
```

```
  TITLE       Methods for large scale production of recombinant dna-derived tpa
              or k2s molecules
  JOURNAL     Patent: WO 0240650-A 6 23-MAY-2002;
              BOEHRINGER INGELHEIM INT (DE)
FEATURES      Location/Qualifiers
     source            1..66
                       /organism="Escherichia coli"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:562"
BASE COUNT      14 a      18 c      21 g      13 t
ORIGIN

Query Match            100.0%;  Score 66; DB 6; Length 66;
Best Local Similarity  100.0%;  Pred. No. 1.4e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
        ||||||
Db   61 GCGGCC 66


RESULT 4
HSIGKLC28
LOCUS         HSIGKLC28             705 bp    mRNA      linear   SYN 29-MAR-2001
DEFINITION    Synthetic construct including Homo sapiens immunoglobulin kappa
              chain (clone: 28).
ACCESSION     X95748
VERSION       X95748.1  GI:1514578
KEYWORDS      constant region; immunoglobulin; kappa light chain.
SOURCE        synthetic construct
  ORGANISM    synthetic construct
              artificial sequences.
REFERENCE     1
  AUTHORS     Steinberger,P., Kraft,D. and Valenta,R.
  TITLE       Construction of a combinatorial IgE library from an allergic
              patient. Isolation and characterization of human IgE Fabs with
              specificity for the major timothy grass pollen allergen, Phl p 5
  JOURNAL     J. Biol. Chem. 271 (18), 10967-10972 (1996)
  MEDLINE     96210038
  PUBMED      8631916
REFERENCE     2  (bases 1 to 705)
  AUTHORS     Valenta,R.L.S.
  TITLE       Direct Submission
  JOURNAL     Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
              Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
              1090 Vienna, AUSTRIA
FEATURES      Location/Qualifiers
     source            1..705
                       /organism="synthetic construct"
                       /mol_type="mRNA"
                       /db_xref="taxon:32630"
                       /clone="28"
                       /cell_type="B-cell"
                       /clone_lib="pcomb3H"
                       /focus
     source            1..66
                       /organism="Escherichia coli"
                       /mol_type="mRNA"
                       /db_xref="taxon:562"
     source            67..705
                       /organism="Homo sapiens"
                       /mol_type="mRNA"
                       /db_xref="taxon:9606"
     gene              1..705
                       /gene="OMP-A"
     CDS               1..705
                       /gene="OMP-A"
                       /codon_start=1
                       /transl_table=11
```

```
AUTHORS   Steinberger,P., Kraft,D., and Valenta,R.
TITLE     Construction of a combinatorial IgE library from an allergic
          patient. Isolation and characterization of human IgE Fabs with
          specificity for the major timothy grass pollen allergen, Phl p 5
JOURNAL   J. Biol. Chem. 271 (18), 10967-10972 (1996)
MEDLINE   96210038
PUBMED    8631916
REFERENCE 2 (bases 1 to 705)
AUTHORS   Valenta,R.L.S.
TITLE     Direct Submission
JOURNAL   Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
          Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
          1090 Vienna, AUSTRIA
FEATURES             Location/Qualifiers
     source          1..705
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /focus
     source          1..66
                     /organism="Escherichia coli"
                     /mol_type="mRNA"
                     /db_xref="taxon:562"
     source          67..705
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="5"
                     /cell_type="B-cell"
                     /clone_lib="pcomb3H"
     gene            1..705
                     /gene="OMP-A"
     CDS             1..705
                     /gene="OMP-A"
                     /codon_start=1
                     /transl_table=11
                     /product="immunoglobulin kappa light chain"
                     /protein_id="CAA65058.1"
                     /db_xref="GI:1514581"
                     /db_xref="REMTREMBL:CAA65058"
                     /translation="MKKTAIAIAVALAGFATVAQAAELTQSPSSVSASVGDRVTITCR
                     ASQGISSWLAWYQQKPGKAPKLLIYSASSLQSGVPSRFSGSGSGTDFSLTISSLQPED
                     SATYYCQQANSFPYTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY
                     PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKLYACEVTH
                     QGLSSPVTKSFNRGEC"
     mRNA            1..66
                     /gene="OMP-A"
                     /note="leaderpeptide gene"
     misc_feature    67..129
                     /gene="OMP-A"
                     /note="FR1"
     primer_bind     67..84
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
     misc_feature    67..72
                     /gene="OMP-A"
                     /note="SAC 1 restriction site"
     misc_feature    130..162
                     /gene="OMP-A"
                     /note="CDR1"
     misc_feature    163..207
                     /gene="OMP-A"
                     /note="FR2"
     misc_feature    208..228
                     /gene="OMP-A"
                     /note="CDR2"
     misc_feature    229..324
                     /gene="OMP-A"
                     /note="FR3"
     misc_feature    325..351
                     /gene="OMP-A"
                     /note="CDR3"
     misc_feature    352..384
```

```
                     /product="immunoglobulin kappa light chain"
                     /protein_id="CAA65059.1"
                     /db_xref="GI:1514579"
                     /db_xref="REMTREMBL:CAA65059"
                     /translation="MKKTAIAIAVALAGFATVAQAAELTQSPSSLSASVGDRVTITCR
                     ASQSILGYLNWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTDFTLTITSLQPDD
                     FATYYCQQSYITPRTFGQGTKVEKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY
                     PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTH
                     QGLSSPVTKFNFRGEC"
     mRNA            1..66
                     /gene="OMP-A"
                     /note="leaderpeptide gene"
     misc_feature    67..129
                     /gene="OMP-A"
                     /note="FR1"
     primer_bind     67..84
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
     misc_feature    67..72
                     /gene="OMP-A"
                     /note="SAC 1 restriction site"
     misc_feature    130..162
                     /gene="OMP-A"
                     /note="CDR1"
     misc_feature    163..207
                     /gene="OMP-A"
                     /note="FR2"
     misc_feature    208..228
                     /gene="OMP-A"
                     /note="CDR2"
     misc_feature    229..324
                     /gene="OMP-A"
                     /note="FR3"
     misc_feature    325..351
                     /gene="OMP-A"
                     /note="CDR3"
     misc_feature    352..384
                     /gene="OMP-A"
                     /note="FR4"
     C_region        385..702
                     /gene="OMP-A"
                     /note="kappa constant region"
     primer_bind     661..705
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
BASE COUNT      184 a    201 c    172 g    148 t
ORIGIN

Query Match     100.0%;  Score 66; DB 12; Length 705;
Best Local Similarity  100.0%;  Pred. No. 1.6e-11;
Matches   66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60

Qy  61 GCGGCC 66
        ||||||
Db  61 GCGGCC 66

RESULT 5
HSIGKLC5
LOCUS       HSIGKLC5        705 bp    mRNA      linear   SYN 29-MAR-2001
DEFINITION  Synthetic construct including Homo sapiens immunoglobulin kappa
            chain (clone: 5).
ACCESSION   X95747
VERSION     X95747.1  GI:1514580
KEYWORDS    constant region; immunoglobulin; kappa light chain.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
```

```
                    /gene="OMP-A"                                                          /protein_id="CAA65060.1"
C_region            /note="FR4"                                                            /db_xref="GI:1514577"
                    385..702                                                                /db_xref="REMTREMBL:CAA65060"
                    /gene="OMP-A"                                                           /translation="MKKTAIAIAVALAGFATVAQAAELVMTQSPSSLSASVGDRVTIT
                    /note="kappa constant region"                                          CRASHSISNYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQP
primer_bind         661..705                                                               EDPANYYCQESFSPSGTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNN
                    /gene="OMP-A"                                                          FYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKLYACEV
                    /note="PCR-primer sequence"                                            THQGLSSPVTKSFNRGEC"
BASE COUNT    183 a   191 c   181 g   150 t                                     mRNA      1..66
ORIGIN                                                                                     /gene="OMP-A"
                                                                                           /note="leaderpeptide gene"
Query Match           100.0%;  Score 66; DB 12; Length 705;                     misc_feature  67..135
Best Local Similarity 100.0%;  Pred. No. 1.6e-11;                                          /gene="OMP-A"
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;                       /note="FR1"
                                                                                primer_bind  67..84
Qy     1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60                     /gene="OMP-A"
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                       /note="PCR-primer sequence"
Db     1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60          misc_feature  67..72
                                                                                           /gene="OMP-A"
Qy    61 GCGGCC 66                                                                         /note="SAC 1 restriction site"
         ||||||                                                                 misc_feature  136..168
Db    61 GCGGCC 66                                                                         /gene="OMP-A"
                                                                                           /note="CDR1"
RESULT 6                                                                        misc_feature  169..213
HSIGKLC14                                                                                  /gene="OMP-A"
LOCUS       HSIGKLC14           711 bp   mRNA   linear   SYN 29-MAR-2001                    /note="FR2"
DEFINITION  Synthetic construct including Homo sapiens immunoglobulin kappa     misc_feature  214..234
            chain (clone: 14).                                                             /gene="OMP-A"
ACCESSION   X95749                                                                         /note="CDR2"
VERSION     X95749.1  GI:1514576                                                misc_feature  235..330
KEYWORDS    constant region; immunoglobulin; kappa light chain.                            /gene="OMP-A"
SOURCE      synthetic construct                                                            /note="FR3"
ORGANISM    synthetic construct                                                 misc_feature  331..357
            artificial sequences.                                                          /gene="OMP-A"
REFERENCE   1                                                                              /note="CDR3"
AUTHORS     Steinberger,P., Kraft,D. and Valenta,R.                             misc_feature  358..390
TITLE       Construction of a combinatorial IgE library from an allergic                   /gene="OMP-A"
            patient. Isolation and characterization of human IgE Fabs with                 /note="FR4"
            specificity for the major timothy grass pollen allergen, Phl p 5   C_region   391..708
JOURNAL     J. Biol. Chem. 271 (18), 10967-10972 (1996)                                    /gene="OMP-A"
MEDLINE     96210038                                                                       /note="kappa constant region"
PUBMED      8631916                                                             primer_bind  667..708
REFERENCE   2  (bases 1 to 711)                                                            /gene="OMP-A"
AUTHORS     Valenta,R.L.S.                                                                 /note="PCR-primer sequence"
TITLE       Direct Submission                                                  BASE COUNT    187 a   194 c   179 g   151 t
JOURNAL     Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &     ORIGIN
            Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
            1090 Vienna, AUSTRIA                                               Query Match           100.0%;  Score 66; DB 12; Length 711;
FEATURES    Location/Qualifiers                                                Best Local Similarity 100.0%;  Pred. No. 1.6e-11;
source      1..711                                                             Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
            /organism="synthetic construct"
            /mol_type="mRNA"                                                   Qy     1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
            /db_xref="taxon:32630"                                                     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            /focus                                                             Db     1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
source      1..66
            /organism="Escherichia coli"                                       Qy    61 GCGGCC 66
            /mol_type="mRNA"                                                           ||||||
            /db_xref="taxon:562"                                               Db    61 GCGGCC 66
source      67..711
            /organism="Homo sapiens"                                           RESULT 7
            /mol_type="mRNA"                                                   HSIGKLG31
            /db_xref="taxon:9606"                                              LOCUS       HSIGKLG31           711 bp   mRNA   linear   SYN 29-MAR-2001
            /clone="14"                                                        DEFINITION  Synthetic construct including Homo sapiens immunoglobulin kappa
            /cell_type="B-cell"                                                            chain (clone: 31).
            /clone_lib="pcomb3H"                                               ACCESSION   X95750
gene        1..711                                                             VERSION     X95750.1  GI:1514582
            /gene="OMP-A"                                                      KEYWORDS    constant region; immunoglobulin; kappa light chain.
CDS         1..711                                                             SOURCE      synthetic construct
            /gene="OMP-A"                                                      ORGANISM    synthetic construct
            /codon_start=1                                                                 artificial sequences.
            /transl_table=11                                                   REFERENCE   1
            /product="immunoglobulin kappa light chain"                        AUTHORS     Steinberger,P., Kraft,D. and Valenta,R.
```

TITLE   Construction of a combinatorial IgE library from an allergic
             patient. isolation and characterization of human IgE Fabs with
             specificity for the major timothy grass pollen allergen, Phl p 5
JOURNAL   J. Biol. Chem. 271 (18), 10967-10972 (1996)
MEDLINE   9210038
PUBMED   8631916
REFERENCE   2 (bases 1 to 711)
AUTHORS   Valenta,R.L.S.
TITLE   Direct Submission
JOURNAL   Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
             Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
             1090 Vienna, AUSTRIA
FEATURES   Location/Qualifiers
     source   1..711
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /focus
     source   1..66
                     /organism="Escherichia coli"
                     /mol_type="mRNA"
                     /db_xref="taxon:562"
     source   67..711
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="31"
                     /cell_type="B-cell"
                     /clone_lib="pcomb3H"
     gene   1..711
                     /gene="OMP-A"
     CDS   1..711
                     /gene="OMP-A"
                     /codon_start=1
                     /transl_table=11
                     /product="immunoglobulin kappa light chain"
                     /protein_id="CAA65061.1"
                     /db_xref="GI:1514583"
                     /db_xref="REMTREMBL:CAA65061"
                     /translation="MKKTAIAIAVALAGFATVAQAAAELVMTQSPSSLSAPVGDRVTIT
                     CRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQP
                     EDFATYCQQSHSTPYTFGQGTNLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNN
                     FYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEV
                     THQGLSSPVTKSFNRGEC"
     mRNA   1..66
                     /gene="OMP-A"
                     /note="leaderpeptide gene"
     misc_feature   67..135
                     /gene="OMP-A"
                     /note="FR1"
     primer_bind   67..84
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
     misc_feature   67..72
                     /gene="OMP-A"
                     /note="SAC 1 restriction site"
     misc_feature   136..168
                     /gene="OMP-A"
                     /note="CDR1"
     misc_feature   169..213
                     /gene="OMP-A"
                     /note="FR2"
     misc_feature   214..234
                     /gene="OMP-A"
                     /note="CDR2"
     misc_feature   235..330
                     /gene="OMP-A"
                     /note="FR3"
     misc_feature   331..357
                     /gene="OMP-A"
                     /note="CDR3"
     misc_feature   358..390
                     /gene="OMP-A"

     C_region   /note="FR4"
                     391..708
                     /gene="OMP-A"
                     /note="kappa constant region"
     primer_bind   667..708
                     /gene="OMP-A"
                     /note="PCR-primer sequence"
BASE COUNT   188 a   194 c   179 g   150 t
ORIGIN
Query Match   100.0%; Score 66; DB 12; Length 711;
Best Local Similarity   100.0%; Pred. No. 1.6e-11;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
               ||||||
Db   61 GCGGCC 66

RESULT 8
AX431268
LOCUS   AX431268   1128 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 2 from Patent WO0240650.
ACCESSION   AX431268
VERSION   AX431268.1 GI:21656150
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 2 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
     source   1..1128
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT   238 a   332 c   332 g   226 t
ORIGIN
Query Match   100.0%; Score 66; DB 6; Length 1128;
Best Local Similarity   100.0%; Pred. No. 1.6e-11;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACACAGCTATCGCGATTGCAGTGGCACTGGCGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
               ||||||
Db   61 GCGGCC 66

RESULT 9
AX431271
LOCUS   AX431271   1128 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 5 from Patent WO0240650.
ACCESSION   AX431271
VERSION   AX431271.1 GI:21656153
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.

```
TITLE       Methods for large scale production of recombinant dna-derived tpa
            or K2s molecules
JOURNAL     Patent: WO 0240650-A 5 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES
   source   1..1128
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT     238 a    332 c    332 g    226 t
ORIGIN

Query Match    100.0%;  Score 66;  DB 6;  Length 1128;
Best Local Similarity  100.0%;  Pred. No. 1.6e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66


RESULT 10
AF268280
LOCUS       AF268280               3394 bp    DNA     circular SYN 03-OCT-2000
DEFINITION  Phagemid cloning vector pComb3H, complete sequence.
ACCESSION   AF268280
VERSION     AF268280.1  GI:10505048
KEYWORDS    .
SOURCE      Phagemid cloning vector pComb3H
  ORGANISM  Phagemid cloning vector pComb3H
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 3394)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  JOURNAL   (in) PHAGE DISPLAY, A LABORATORY MANUAL, Cold Spring Harbor
            Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press
REFERENCE   2  (bases 1 to 3394)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,
            Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA
            92037, USA
FEATURES
   source   1..3394
            /organism="Phagemid cloning vector pComb3H"
            /mol_type="genomic DNA"
            /db_xref="taxon:137785"
BASE COUNT     862 a    783 c    876 g    873 t
ORIGIN

Query Match    100.0%;  Score 66;  DB 12;  Length 3394;
Best Local Similarity  100.0%;  Pred. No. 1.7e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 281

Qy   61 GCGGCC 66
          ||||||
Db  282 GCGGCC 287


RESULT 11
AR241678
LOCUS       AR241678               3450 bp    DNA     linear      PAT 20-DEC-2002
DEFINITION  Sequence 20 from patent US 6472147.
ACCESSION   AR241678
VERSION     AR241678.1  GI:27287480
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3450)
  AUTHORS   Janda,K.D., Wirsching,P., Lerner,R.A. and Gao,C.
  TITLE     Methods for display of heterodimeric proteins on filamentous phage
            using pVII and pIX, compositions, vectors and combinatorial
            libraries
  JOURNAL   Patent: US 6472147-A 20 29-OCT-2002;
FEATURES
   source   1..3450
            /organism="unknown"
BASE COUNT     873 a    799 c    891 g    887 t
ORIGIN

Query Match    100.0%;  Score 66;  DB 6;  Length 3450;
Best Local Similarity  100.0%;  Pred. No. 1.7e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2667 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 2726

Qy   61 GCGGCC 66
          ||||||
Db 2727 GCGGCC 2732


RESULT 12
AF268281
LOCUS       AF268281               3758 bp    DNA     circular SYN 03-OCT-2000
DEFINITION  Phagemid cloning vector pComb3X, complete sequence.
ACCESSION   AF268281
VERSION     AF268281.1  GI:10505049
KEYWORDS    .
SOURCE      Phagemid cloning vector pComb3X
  ORGANISM  Phagemid cloning vector pComb3X
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 3758)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  JOURNAL   (in) PHAGE DISPLAY, A LABORATORY MANUAL, Cold Spring Harbor
            Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press
REFERENCE   2  (bases 1 to 3758)
  AUTHORS   Rader,C. and Barbas,C.F. III.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,
            Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA
            92037, USA
FEATURES
   source   1..3758
            /organism="Phagemid cloning vector pComb3X"
            /mol_type="genomic DNA"
            /db_xref="taxon:137786"
BASE COUNT     939 a    896 c    959 g    964 t
ORIGIN

Query Match    100.0%;  Score 66;  DB 12;  Length 3758;
Best Local Similarity  100.0%;  Pred. No. 1.7e-11;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGGTTTCGCTACCGTGGCCCAG 281

Qy   61 GCGGCC 66
          ||||||
Db  282 GCGGCC 287


RESULT 13
AX591634
LOCUS       AX591634               5149 bp    DNA     linear      PAT 27-JAN-2003
```

```
DEFINITION  Sequence 8 from Patent WO0246434.
ACCESSION   AX591634
VERSION     AX591634.1  GI:27950030
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bowdish,K.S. and Barbas-Frederickson,S.
  TITLE     Novel plasmid vectors
  JOURNAL   Patent: WO 0246434-A 8 13-JUN-2002;
            ALEXION PHARMACEUTICALS, INC. (US)
FEATURES             Location/Qualifiers
     source          1..5149
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="vector"
BASE COUNT    1279 a   1244 c   1372 g   1254 t
ORIGIN

Query Match             100.0%;  Score 66;  DB 6;  Length 5149;
Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676


RESULT 14
AX600065
LOCUS       AX600065    5149 bp    DNA    linear   PAT 14-FEB-2003
DEFINITION  Sequence 60 from Patent WO0246238.
ACCESSION   AX600065
VERSION     AX600065.1  GI:28400138
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bowdish,K.S., Barbas-Frederickson,S. and Renshaw,M.
  TITLE     Rationally designed antibodies
  JOURNAL   Patent: WO 0246238-A 60 13-JUN-2002;
            ALEXION PHARMACEUTICALS, INC. (US)
FEATURES             Location/Qualifiers
     source          1..5149
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="vector"
BASE COUNT    1279 a   1244 c   1372 g   1254 t
ORIGIN

Query Match             100.0%;  Score 66;  DB 6;  Length 5149;
Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676


RESULT 15
AX554424
LOCUS       AX554424    5683 bp    DNA    linear   PAT 27-NOV-2002
DEFINITION  Sequence 12 from Patent WO0246436.
ACCESSION   AX554424
VERSION     AX554424.1  GI:25898200
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bowdish,K.S., Barbas-Frederickson,S., Wild,M. and Mcwhirter,J.
  TITLE     Novel plasmid vectors
  JOURNAL   Patent: WO 0246436-A 12 13-JUN-2002;
            ALEXION PHARMACEUTICALS, INC. (US)
FEATURES             Location/Qualifiers
     source          1..5683
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid"
BASE COUNT    1337 a   1492 c   1543 g   1311 t
ORIGIN

Query Match             100.0%;  Score 66;  DB 6;  Length 5683;
Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy     61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676


Search completed: August 19, 2003, 04:46:31
Job time : 281.124 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 19:51:53 ; Search time 23.2513 Seconds
(without alignments)
7662.485 Million cell updates/sec

Title: US-09-987-455-3
Perfect score: 66
Sequence: 1 atgaaaaagacagctatcgc.........ctaccgtggcccaggcggcc 66

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03;*
1: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	66	24	AAD40615	Escherichia coli o
2	66	100.0	66	24	AAD38987	Escherichia coli O
3	66	100.0	1128	24	AAD40614	OmpA-K2S fusion pr
4	66	100.0	1574	24	AA257599	Erythrocyte and He
5	66	100.0	3450	22	AAC66993	Vector pCGMT-1b.
6	66	100.0	4691	16	AAQ92546	pComb3 expression
7	66	100.0	5149	24	ABO73372	Plasmid pRL8 nucle
8	66	100.0	5149	24	ABN84077	Phagemid vector pR
9	66	100.0	5683	24	AAL41116	Plasmid pRL5 nucle
10	66	100.0	6122	24	AAL41112	Plasmid pRL5 CAT n
11	66	100.0	6122	24	AAL41115	Plasmid pRL5 bsi-C
12	66	100.0	6122	24	ABN84078	Plasmid pRL5-CAT.
13	66	100.0	6125	24	AAL41114	Plasmid pRL5 CAT-A
14	66	100.0	6166	16	AAQ92547	Expression vector.
15	60.2	91.2	73	12	AAQ10569	Omp A signal pepti
16	60.2	91.2	103	22	AAS44154	Neisseria meningit
17	60.2	91.2	249	11	AAQ03521	TAC recombinant DN
18	60.2	91.2	548	18	AAT91586	Growth factor LHL
19	60.2	91.2	548	20	AAX34590	DNA encoding a LHL
20	60.2	91.2	652	9	AAN81344	sequence encoding
21	60.2	91.2	2814	14	AAQ41731	PE binding/translo
22	60.2	91.2	2814	14	AAQ38414	PE binding and tra
23	60.2	91.2	6477	11	AAQ02030	Plasmid pinf 4-49.
24	59.6	90.3	73	21	AAA12579	Synthetic DNA enco
25	59.6	90.3	415	17	AAT15982	pSKAN8 fragment #1
26	59.6	90.3	470	18	AAT91590	Kappa light chain
27	59.6	90.3	470	20	AAX34594	DNA encoding a hum
28	59.6	90.3	599	18	AAT91588	Growth factor LHL.
29	59.6	90.3	599	20	AAX34593	DNA encoding a for
30	59.6	90.3	759	19	AAV53721	The ORF of the clo
31	59.6	90.3	770	19	AAV53720	Cloned duplicate o
32	59.6	90.3	793	20	AAX24101	Plasmid pBBP22 DNA
33	59.6	90.3	793	22	AAF25703	Bilin binding-prot
34	59.6	90.3	803	12	AAQ14831	Qm212 single chain
35	59.6	90.3	932	22	AAQ91151	Expression cassett
36	59.6	90.3	1031	18	AAT91589	Growth factor TLHL
37	59.6	90.3	1031	20	AAX34592	DNA encoding TLHL.
38	59.6	90.3	1219	20	AAF25695	Plasmid pBBP20 DNA
39	59.6	90.3	1219	22	AAF25708	Bilin binding-prot
40	59.6	90.3	1380	20	AAX24100	Plasmid pBBP21 DNA
41	59.6	90.3	1380	22	AAF25710	Bilin binding-prot
42	59.6	90.3	1479	20	AAX34596	DNA encoding a ccM
43	59.6	90.3	1490	18	AAT91587	Growth factor CATA
44	59.6	90.3	1490	20	AAX34591	DNA encoding CATAB

ALIGNMENTS

RESULT 1
AAD40615
ID   AAD40615 standard; DNA; 66 BP.
XX
AC   AAD40615;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Escherichia coli ompA peptide encoding DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..66
FT   /*tag= a
FT   /product= "OmpA peptide"
FT   /note= "No stop codon"
FT   /partial
XX
PN   WO20020650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.

```
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25035.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 8; Page 30; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA fragment encoding Escherichia coli OmpA peptide.
XX
SQ   Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 66;
Best Local Similarity   100.0%;  Pred. No. 2.1e-14;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCAG 60

Qy         61 GCGGCC 66
              ||||||
Db         61 GCGGCC 66

RESULT 2
AAD38987
ID   AAD38987 standard; DNA; 66 BP.
XX
AC   AAD38987;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Escherichia coli OmpA DNA.
XX
KW   Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;
KW   serine protease; ds.
XX
OS   Escherichia coli.
XX
PN   WO200240696-A2.
XX
PD   23-MAY-2002.
XX
PF   08-NOV-2001; 2001WO-EP12920.
XX
PR   14-NOV-2000; 2000GB-0027782.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
```

```
PT   Producing recombinant DNA-derived kringle 2 plus serine protease,
PT   comprises using a prokaryotic cell expressing a vector having a DNA
PT   coding for a heterologous protein operably linked to a DNA coding for
PT   the signal peptide OmpA -
XX
PS   Claim 9; Page 23; 52pp; English.
XX
CC   The invention relates to a method for producing recombinant DNA-derived
CC   heterologous protein in prokaryotic cells, where the heterologous protein
CC   is secreted extracellularly as an active and correctly folded protein and
CC   the prokaryotic cell contains and expresses a vector comprising the DNA
CC   coding for the heterologous protein operably linked to the DNA coding
CC   for the signal peptide OmpA or its functional derivative. The method is
CC   useful for commercial large-scale production of heterologous proteins,
CC   e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is
CC   generally applicable in the expression of several different proteins and
CC   polypeptides which do not require mammalian glycosylation in prokaryotic
CC   host cells. The method may also be used to obtain DNA sequences of a
CC   protein of interest to be expressed from databases and cloned for use.
CC   The present sequence is Escherichia coli OmpA DNA.
XX
SQ   Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 66;
Best Local Similarity   100.0%;  Pred. No. 2.1e-14;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCAG 60

Qy         61 GCGGCC 66
              ||||||
Db         61 GCGGCC 66

RESULT 3
AAD40614
ID   AAD40614 standard; DNA; 1128 BP.
XX
AC   AAD40614;
XX
DT   30-OCT-2002  (first entry)
XX
DE   OmpA-K2S fusion protein encoding DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; fusion protein; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1128
FT                   /*tag= a
FT                   /product= "OmpA-K2S fusion protein"
FT                   /transl_except= (pos:1126..1128, aa:Pro-Gly)
FT                   /note="No stop codon"
FT                   /partial
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
```

DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25034.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 7; Page 30; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA encoding a fusion protein comprising OmpA and K2S protein.
XX
SQ   Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;

Query Match   100.0%;   Score 66; DB 24; Length 1128;
Best Local Similarity 100.0%; Pred. No. 4.3e-14;
Matches 66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
         ||||||
Db   61 GCGGCC 66

RESULT 4
AAZ57599
ID   AAZ57599 standard; DNA; 1574 BP.
XX
AC   AAZ57599;
XX
DT   19-APR-2000 (first entry)
XX
DE   Erythrocyte and Hepatitis B dual-specific antibody DNA sequence.
XX
KW   Erythrocyte; Hepatitis B; dual-specific antibody; surface antigen; ds.
XX
OS   Unidentified.
XX
PN   CN1232039-A.
XX
PD   20-OCT-1999.
XX
PF   02-APR-1999; 99CN-0103517.
XX
PR   02-APR-1999; 99CN-0103517.
XX
PA   (NAVA-) NAVAL GEN HOSPITAL PLA.
XX
PI   Chen Y, Wang Y;
XX
DR   WPI; 2000-098467/09.
XX
PT   Genetic engineering double specific antibody and its use -
XX
PS   Claim 1-2; 6pp; Chinese.
XX
CC   The present invention describes a dual-specific antibody composed of
CC   the surface antigens to resist against erythrocyte and hepatitis B,

CC   which can be used as the test reagent to detect the surface antigen
CC   of hepatitis B in blood. The dual-specific antibody is prepared by
CC   recombination technique in gene engineering. The recombination technique
CC   includes reforming the surface antigen to resist against erythrocyte
CC   and hepatitis B to become hybrid antibody genes by shortening the
CC   joining peptide of single-chain antibody, and assembling them in the
CC   same expression vector. The gene product can be directly extracted from
CC   the supernatant of bacterial culture liquid. Its advantages are low
CC   cost, quick detection and simple operation.
XX
SQ   Sequence 1574 BP; 395 A; 390 C; 419 G; 370 T; 0 other;

Query Match   100.0%;   Score 66; DB 21; Length 1574;
Best Local Similarity 100.0%; Pred. No. 4.6e-14;
Matches 66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 792 ATGAAAAAGACAGCTATCGGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 851

Qy   61 GCGGCC 66
         ||||||
Db 852 GCGGCC 857

RESULT 5
AAC66993
ID   AAC66993 standard; DNA; 3450 BP.
XX
AC   AAC66993;
XX
DT   28-MAR-2001 (first entry)
XX
DE   Vector pCGMT-1b.
XX
KW   Filamentous phage; protein display; pVII; pIX;
KW   combinatorial antibody library; ds.
XX
OS   Synthetic.
XX
PN   WO200071694-A1.
XX
PD   30-NOV-2000.
XX
PF   24-MAY-2000; 2000WO-US14433.
XX
PR   25-MAY-1999; 99US-0318786.
XX
PA   (SCRI) SCRIPPS RES INST.
XX
PI   Janda KD, Wirsching P, Lerner RA, Gao C;
XX
DR   WPI; 2001-032030/04.
XX
PT   Novel filamentous phage encapsulating a genome encoding fusion
PT   polypeptide comprising exogenous polypeptide fused to amino terminus of
PT   pVII and pIX proteins, for constructing diverse heterodimeric
PT   polypeptide array -
XX
PS   Example 1; Page 74-75; 90pp; English.
XX
CC   The present invention describes a filamentous phage encapsulating a
CC   genome encoding a fusion protein. This fusion protein comprises an
CC   exogenous protein fused to the amino terminus of a filamentous phage pVII
CC   or pIX protein. This is useful in the design of proteins for medical,
CC   industrial, environmental and research applications.
XX
SQ   Sequence 3450 BP; 873 A; 799 C; 891 G; 887 T; 0 other;

Query Match   100.0%;   Score 66; DB 22; Length 3450;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2667 ATGAAAAAGACAGCTATCCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2726

Qy   61 GCGGCC 66
          ||||||
Db   2727 GCGGCC 2732

RESULT 6
AAQ92546
ID AAQ92546 standard; DNA; 4691 BP.
XX
AC AAQ92546;
XX
DT 11-MAR-1996 (first entry)
XX
DE pComb3 expression vector.
XX
KW pComb3; phagemid expression vector; bacteriophage; coat protein 3;
KW Gene III; filamentous phage; minor phage coat protein; cPIII; cp3;
KW bacterial membrane; periplasm; E. coli; human; Fab; HIV; gp120;
KW combinatorial Fab library; cassette; Fd/cp3; lacZ promoter/operator;
KW ribosome binding site; RBS; PelB leader; spacer; tether sequence;
KW MT4; pMT4-3; antibody; ss; cyclic.
XX
OS Synthetic.
XX
PN WO9511317-A1.
XX
PD 27-APR-1995.
XX
PF 19-OCT-1994;   94WO-US11907.
XX
PR 19-SEP-1994;   94US-0308841.
PR 19-OCT-1993;   93US-0139409.
PR 26-APR-1994;   94US-0233619.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Burton DR, Lerner RA;
XX
DR WPI; 1995-170235/22.
XX
PT Synthetic human neutralising monoclonal antibodies to human
PT immunodeficiency virus - used for diagnosis and immuno:therapy of
PT HIV-induced disease
XX
PS Example 1; Page 185-188; 249pp; English.
XX
CC This sequence represents the pComb3 phagemid expression vector. This
CC vector has been designed to allow for anchoring of expressed proteins on
CC the bacteriophage coat protein 3. Gene III of filamentous phage encodes
CC the 406 residue minor phage coat protein, cpIII (cp3), which is
CC expressed prior to extrusion in the phage assembly process on a bacterial
CC membrane and accumulates on the inner membrane facing into the periplasm
CC of E. coli. This plasmid was used within the scope of the invention to
CC express various mutagenised human Fab's which comprise heavy and light
CC variable regions which bind to HIV gp120. pComb3 allows for both surface
CC display and soluble forms of the Fabs. The vector was designed for the
CC cloning of combinatorial Fab libraries. pComb consists of a DNA molecule
CC having two cassettes to express one fusion protein, Fd/cp3, and one
CC soluble protein, the light chain. The finished vector comprises,
CC operatively linked 5' to 3', a first cassette consisting of lacZ
CC promoter/operator sequences, a NotI restriction site, a ribosome binding
CC site (RBS), a PelB leader, a spacer region, a cloning region bordered by
CC 5' XhoI and 3' SpeI restriction sites, the tether sequence, the sequence
CC encoding bacteriophage cp3 followed by a stop codon, a NheI restriction
CC site between the two cassettes, and a second lacZ promoter/operator
CC sequence, followed by an expression control RBS, a PelB leader, a spacer
CC region, a cloning region bordered by 5' SacI and 3' XbaI restriction
CC sites, followed by expression control stop sequences and a second NotI
CC restriction site. The pComb3 expression vector forms the basic

CC construct of the MT4 Fab display phagemid expression vector, pMT4-3 (see
CC also AAQ92540), used in the invention for the production of synthetic
CC human Fab antibodies against gp120 of HIV.
XX
SQ Sequence 4691 BP; 1170 A; 1171 C; 1232 G; 1118 T; 0 other;

Query Match   100.0%; Score 66; DB 16; Length 4691;
Best Local Similarity 100.0%; Pred. No. 6.1e-14;
Matches 66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
          ||||||
Db   2671 GCGGCC 2676

RESULT 7
ABQ73372
ID ABQ73372 standard; DNA; 5149 BP.
XX
AC ABQ73372;
XX
DT 01-OCT-2002 (first entry)
XX
DE Plasmid pRL8 nucleotide sequence SEQ ID NO:60.
XX
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobin; antianaemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis;
KW gene; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200246238-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US47656.
XX
PR 05-DEC-2000; 2000US-251448P.
PR 04-MAY-2001; 2001US-288889P.
PR 29-MAY-2001; 2001US-294068P.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Barbas-Frederickson S, Renshaw M;
XX
DR WPI; 2002-566610/60.
XX
PT A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
PT thrombopoetin mimetic -
XX
PS Disclosure; Fig 6A-C; 113pp; English.
XX
CC The present invention describes an immunoglobin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC promegakaryocytes or megakaryocytes, which results in increased platelet
CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (II) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
CC in the exemplification of the present invention.
XX
SQ Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 6.2e-14;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2611 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy 61 GCGGCC 66
       ||||||
Db 2671 GCGGCC 2676

RESULT 8
ABN84077
ID ABN84077 standard; DNA; 5149 BP.
XX
AC ABN84077;
XX
DT 23-SEP-2002 (first entry)
XX
DE Phagemid vector pRL8.
XX
KW Vector; phagemid; pRL8; pComb; dimerisation domain; Jun;
KW leucine zipper; antibody; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT misc_feature 2677..3866
FT /tag= a
FT /note= "light chain stuffer region"
FT misc_feature 3968..4337
FT /tag= b
FT /note= "heavy chain stuffer region"
FT CDS 4338..4583
FT /tag= c
FT /partial
FT /product= "Jun dimerisation domain"
FT /note= "the CDS does not include a start codon"
XX
PN WO20024 6434-A2.
XX
PD 13-JUN-2002.
XX
PF 07-DEC-2001; 2001WO-US46514.
XX
PR 08-DEC-2000; 2000US-254410P.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Barbas-Frederickson S;
XX
DR WPI; 2002-537568/57.
DR P-PSDB; ABB79463.
XX
PT Novel pComb family phagemid comprising a nucleic acid encoding a
PT dimerization domain, useful for producing antibody fragments which
PT require dimerization in order to crosslink for activation of target
PT receptors -
XX

PS Disclosure; Fig 8A-C; 17pp; English.
XX
CC The present sequence is of plasmid pLR8, a pComb vector family
CC member. pLR8 was constructed by insertion of a PCR fragment
CC containing a flexible linker (murine kappa hinge region) followed
CC by a Jun leucine zipper dimerisation domain into a pRL4 backbone
CC that had been treated to remove the stuffer region. The resulting
CC construct was then cut with restriction endonucleases to replace
CC scFv with a pRL4 stuffer region for the light and heavy chain. The
CC vector is useful in the production of biologically active proteins
CC such as antibody fragments which require dimerisation for their
CC biological activity. It can also be used in the production and
CC screening of libraries made in accordance with phage display
CC technology. Homo-dimerisation of single chain antibodies is
CC accomplished by the presence of the Jun leucine zipper domain,
CC which is responsible for protein-protein interactions. Subcloning
CC to pRL8 individually or en masse following FACS sorting or panning
CC allows expression, e.g., of dimeric soluble Fabs.
XX
SQ Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 6.2e-14;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2611 ATGAAAAAGACAGCTATCGCGATTGCCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy 61 GCGGCC 66
       ||||||
Db 2671 GCGGCC 2676

RESULT 9
AAL41116
ID AAL41116 standard; DNA; 5683 BP.
XX
AC AAL41116;
XX
DT 16-OCT-2002 (first entry)
XX
DE Plasmid pRL5 nucleic acid sequence.
XX
KW Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 591..1451
FT /tag= a
FT /product= "Beta lactamase"
FT CDS 2611..2676
FT /tag= b
FT /product= "Omp A leader amino acid sequence"
FT CDS 3508..3828
FT /tag= c
FT /product= "Kappa constant region amino acid sequence"
FT CDS 3859..3924
FT /tag= d
FT /product= "pel B leader amino acid sequence"
FT CDS 4740..5651
FT /tag= e
FT /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT sequence"
XX
PN WO20024 6436-A2.
XX
PD 13-JUN-2002.
XX
PF 07-DEC-2001; 2001WO-US46516.
XX

```
XX
PR   08-DEC-2000; 2000US-254411P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;
XX
DR   WPI; 2002-537570/57.
DR   P-PSDB; AAO22536, AAO22537, AAO22538, AAO22539, AAO22542.
XX
PT   Novel plasmid useful in cloning and expression of foreign genetic
PT   information -
XX
PS   Claim 1; Fig 4A-E; 39pp; English.
XX
CC   The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC   6125, or 5683 base pairs fully defined in the specification. The
CC   invention more specifically relates to novel vectors capable of
CC   replication and expression of foreign genetic information in bacteria,
CC   such as, for example, cyanobacterium and E. coli. The new vectors have
CC   been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC   new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC   polynucleotide sequence represents the plasmid pRL5 nucleic acid sequence
CC   of the invention.
XX
SQ   Sequence 5683 BP; 1337 A; 1495 C; 1540 G; 1311 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 5683;
Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy         61 GCGGCC 66
              ||||||
Db       2671 GCGGCC 2676

RESULT 10
AAL41112
ID   AAL41112 standard; DNA; 6122 BP.
XX
AC   AAL41112;
XX
DT   16-OCT-2002 (first entry)
XX
DE   Plasmid pRL5 CAT nucleic acid sequence.
XX
KW   Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW   cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW   pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1013..1672
FT                   /*tag= a
FT                   /product= "CAT amino acid sequence"
FT   CDS             3050..3115
FT                   /*tag= b
FT                   /product= "Omp A leader amino acid sequence"
FT   CDS             3953..4267
FT                   /*tag= c
FT                   /product= "Kappa constant region amino acid sequence"
FT   CDS             4298..4363
FT                   /*tag= d
FT                   /product= "pel B leader amino acid sequence"
FT   CDS             5179..6090
FT                   /*tag= e
FT                   /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT                   sequence"
```

```
XX
PN   WO200246436-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46516.
XX
PR   08-DEC-2000; 2000US-254411P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;
XX
DR   WPI; 2002-537570/57.
DR   P-PSDB; AAO22535, AAO22536, AAO22537, AAO22538, AAO22539.
XX
PT   Novel plasmid useful in cloning and expression of foreign genetic
PT   information -
XX
PS   Claim 1; Fig 7A-E; 39pp; English.
XX
CC   The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC   6125, or 5683 base pairs fully defined in the specification. The
CC   invention more specifically relates to novel vectors capable of
CC   replication and expression of foreign genetic information in bacteria,
CC   such as, for example, cyanobacterium and E. coli. The new vectors have
CC   been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC   new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC   polynucleotide sequence represents the plasmid pRL5-CAT nucleic acid
CC   sequence of the invention.
XX
SQ   Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;

Query Match             100.0%;  Score 66;  DB 24;  Length 6122;
Best Local Similarity   100.0%;  Pred. No. 6.5e-14;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy         61 GCGGCC 66
              ||||||
Db       3110 GCGGCC 3115

RESULT 11
AAL41115
ID   AAL41115 standard; DNA; 6122 BP.
XX
AC   AAL41115;
XX
DT   16-OCT-2002 (first entry)
XX
DE   Plasmid pRL5 bsi-CAT nucleic acid sequence.
XX
KW   Plasmid; vector; replication; expression; foreign genetic; bacteria;
KW   cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW   pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1013..1672
FT                   /*tag= a
FT                   /product= "CAT amino acid sequence"
FT   CDS             3050..3115
FT                   /*tag= b
FT                   /product= "Omp A leader amino acid sequence"
FT   CDS             3947..4267
FT                   /*tag= c
FT                   /product= "Kappa constant region amino acid sequence"
FT   CDS             4298..4363
```

```
FT                      /*tag= d
FT                      /product= "pel B leader amino acid sequence"
FT                      5179..6090
FT                      /*tag= e
FT   CDS                /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT                      sequence"
XX
PN   WO200246436-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46516.
XX
PR   08-DEC-2000; 2000US-254411P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;
XX
DR   WPI; 2002-537570/57.
DR   P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22541.
XX
PT   Novel plasmid useful in cloning and expression of foreign genetic
PT   information.
XX
PS   Disclosure; Fig 11A-E; 39pp; English.
XX
CC   The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC   6125, or 5683 base pairs fully defined in the specification. The
CC   invention more specifically relates to novel vectors capable of
CC   replication and expression of foreign genetic information in bacteria,
CC   such as, for example, cyanobacterium and E. coli. The new vectors have
CC   been designed to overcome certain drawbacks of the pComb3x plasmid. These
CC   new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC   polynucleotide sequence represents the plasmid pRL5 bsi-CAT nucleic acid
CC   sequence of the invention.
XX
SQ   Sequence 6122 BP; 1447 A; 1571 C; 1641 G; 1463 T; 0 other;

Query Match          100.0%;  Score 66;  DB 24;  Length 6122;
Best Local Similarity   100.0%;  Pred. No. 6.5e-14;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGTTTCGCTACCGTGGCCCAG 3109

Qy     61 GCGGCC 66
          ||||||
Db   3110 GCGGCC 3115


RESULT 12
ABN84078
ID   ABN84078 standard; DNA; 6122 BP.
XX
AC   ABN84078;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Plasmid pRL5-CAT.
XX
KW   Plasmid pTL5-CAT; vector; antibody; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             591..800
FT                   /*tag= a
FT                   /partial
FT                   /product= "Amp frag"
FT                   /note= "the CDS does not include a stop codon"
FT   CDS             1013..1672
```

```
FT   misc_feature    /*tag= b
FT                   /product= "Chloramphenicol transferase"
FT                   complement (2052..2651)
FT                   /*tag= c
FT                   /note= "ori"
FT   promoter        2831..3046
FT                   /*tag= d
FT                   /note= "lac promoter"
FT   misc_signal     3009..3029
FT                   /*tag= e
FT                   /note= "lac rep site"
FT   RBS             3036..3039
FT                   /*tag= f
FT   CDS             3050..3115
FT                   /*tag= g
FT                   /partial
FT                   /product= "OmpA leader"
FT                   /note= "the CDS does not include a stop codon"
FT   misc_feature    3121..3946
FT                   /*tag= h
FT                   /note= "light chain variable region stuffer"
FT   misc_feature    3951..4269
FT                   /*tag= i
FT                   /note= "Kappa Cns"
FT   CDS             4298..4363
FT                   /*tag= j
FT                   /partial
FT                   /product= "pelB leader"
FT                   /note= "the CDS does not include a stop codon"
FT   misc_feature    4365..5147
FT                   /*tag= k
FT                   /note= "heavy chain variable region stuffer"
FT   CDS             5485..5556
FT                   /*tag= l
FT                   /partial
FT                   /product= "Linker-His tag-HA tag"
FT                   /note= "the CDS does not include a start codon"
FT   CDS             5557..6090
FT                   /*tag= m
FT                   /partial
FT                   /gene= "Gene III"
FT                   /note= "the CDS does not include a start codon"
XX
PN   WO200246435-A2.
XX
PD   13-JUN-2002.
XX
PF   05-DEC-2001; 2001WO-US47452.
XX
PR   05-DEC-2000; 2000US-251440P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS, Barbas-Frederickson S, Lin Y, Renshaw M, Wild M;
PI   McWhirter J;
XX
DR   WPI; 2002-537569/57.
XX
PT   Plasmid for in situ production of genes, comprises two template
PT   annealing sequences, such as downstream primer and upstream collar
PT   sequence and a restriction site located between the annealing sequences
PT   -
XX
PS   Disclosure; Fig 4A-T; 65pp; English.
XX
CC   The present sequence is that of plasmid pRL5-CAT, a derivative of
CC   pComb 3X which has been modified to contain chloramphenicol
CC   resistance. The invention relates to the engineering of plasmids,
CC   of which pRL-CAT is a preferred example, for in situ production of
CC   genes. It was found that nucleic acids encoding a polypeptide
CC   can be directly incorporated into a plasmid by DNA polymerisation
CC   or by reverse transcription of a nucleic acid template. Preferably,
CC   nucleic acids encoding at least a portion of an antibody can be
```

CC   directly incorporated into the plasmid by reverse transcription of
CC   mRNA. The plasmids are engineered to contain 2 template annealing
CC   sequences (see ABN84079-80 and ABN84082-83), i.e. a downstream
CC   primer that anneals to a first portion of a nucleic acid template,
CC   e.g. mRNA encoding at least a portion of an antibody, an upstream
CC   collar sequence that anneals to a second portion of the template,
CC   and at least 1 restriction site located between the 2 template
CC   annealing sequences. A single-stranded DNA plasmid vector is
CC   produced containing a nucleic acid encoding at least a portion of
CC   a polypeptide, e.g. a light chain and/or a heavy chain of an
CC   antibody. This vector can be transformed into a host cell and
CC   amplified.
XX
SQ   Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;

Query Match   100.0%; Score 66; DB 24; Length 6122;
Best Local Similarity 100.0%; Pred. No. 6.5e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 3109

Qy   61 GCGGCC 66
         ||||||
Db 3110 GCGGCC 3115

RESULT 13
AAL41114
ID   AAL41114 standard; DNA; 6125 BP.
XX
AC   AAL41114;
XX
DT   16-OCT-2002 (first entry)
XX
DE   Plasmid pRL5 CAT-Asc nucleic acid sequence.
XX
KW   plasmid; vector; replication; expression; foreign genetic; bacteria;
KW   cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
KW   pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1013..1672
FT   /tag= a
FT   /product= "CAT amino acid sequence"
FT   CDS   3050..3115
FT   /tag= b
FT   /product= "Omp.A leader amino acid sequence"
FT   CDS   3953..4270
FT   /tag= c
FT   /product= "Kappa constant region amino acid sequence"
FT   CDS   4301..4366
FT   /tag= d
FT   /product= "pel B leader amino acid sequence"
FT   CDS   5182..6093
FT   /tag= e
FT   /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT   sequence"
XX
PN   WO20024636-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46516.
XX
PR   08-DEC-2000; 2000US-254411P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
RI   Bowdish KS, Barbas-frederickson S, Wild M, Mcwhirter J;

XX
DR   WPI; 2002-537570/57.
DR   P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22540.
XX
PT   Novel plasmid useful in cloning and expression of foreign genetic
PT   information -
XX
PS   Claim 1; Fig 9A-E; 39pp; English.
XX
CC   The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC   6125, or 5683 base pairs fully defined in the specification. The
CC   invention more specifically relates to novel vectors capable of
CC   replication and expression of foreign genetic information in bacteria,
CC   such as, for example, cyanobacterium and E. coli. The new vectors have
CC   been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC   new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC   polynucleotide sequence represents the plasmid pRL5 CAT-Asc nucleic acid
CC   sequence of the invention.
XX
SQ   Sequence 6125 BP; 1447 A; 1573 C; 1644 G; 1461 T; 0 other;

Query Match   100.0%; Score 66; DB 24; Length 6125;
Best Local Similarity 100.0%; Pred. No. 6.5e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3050 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGGTTTCGCTACCGTGGCCCAG 3109

Qy   61 GCGGCC 66
         ||||||
Db 3110 GCGGCC 3115

RESULT 14
AAQ92547
ID   AAQ92547 standard; DNA; 6166 BP.
XX
AC   AAQ92547;
XX
DT   11-MAR-1996 (first entry)
XX
DE   Expression vector, pPho-TT.
XX
KW   Human; Fab; variable chain; heavy; light; region; VH; VL; HIV; gp120;
KW   3b1; 3b3; 3b4; 3b9; MT4; humanised; monoclonal antibody; MAb;
KW   immunoreaction; neutralisation; passive immunotherapy; tetanus toxin;
KW   alkaline phosphatase; phoA; ss; cyclic.
XX
OS   Synthetic.
XX
PN   WO9511317-A1.
XX
PD   27-APR-1995.
XX
PF   19-OCT-1994;   94WO-US11907.
XX
PR   19-SEP-1994;   94US-0308841.
PR   19-OCT-1993;   93US-0139409.
PR   26-APR-1994;   94US-0233619.
XX
PA   (SCRI) SCRIPPS RES INST.
XX
PI   Barbas CF, Burton DR, Lerner RA;
XX
DR   WPI; 1995-170235/22.
XX
PT   Synthetic human neutralising monoclonal antibodies to human
PT   immunodeficiency virus - used for diagnosis and immunotherapy of
PT   HIV-induced disease
XX
PS   Example 2; Page 193-197; 249pp; English.
XX

CC   This sequence represents the expression vector, pPho-TT which is a
CC   modified version of the phagemid expression vector, pComb3 given in
CC   AAQ92546. pPho-TT provides for the expression of soluble Fabs which are
CC   secreed into the periplasmic space which is regulated from the alkaline
CC   phosphatase (phoA) promoter. This plasmid was used within the scope of
CC   the invention to express various mutagenised human Fab's which comprise
CC   heavy and light variable regions which bind to HIV gp120. pPho-TT
CC   consists of a DNA molecule having two cassettes to express two soluble
CC   proteins a heavy chain and a light chain. The vector comprises,
CC   operatively linked 5' to 3', a first cassette consisting of the phoA
CC   promoter/operator sequences, an EcoRI restriction site, a ribosome
CC   binding site (RBS), an OmpA leader, a SfiI restriction site, a spacer
CC   region, a cloning region bordered by 5' sacI and 3' XbaI restriction
CC   sites, an NcoI restriction site between the two cassettes, and a second
CC   cassette consisting of an expression control RBS, a PelB leader, a human
CC   consensus amino terminus spacer region comprising the sequence EVQLLE,
CC   a cloning region bordered by 5' XhoI and 3' SpeI restriction sites
CC   followed by a SfiI site, expression control stop sequences and a NotI
CC   restriction site. The pPho-TT expression vector contains a light
CC   chain stuffer that is 1200 bp in length and a heavy chain stuffer that
CC   is 300 bp in length. The nucleotide sequences of the heavy and light
CC   chain stuffers encode the heavy and light chain variable domains of a
CC   tetanus toxin-specific Fab.
XX
SQ   Sequence 6166 BP; 1416 A; 1706 C; 1629 G; 1415 T; 0 other;

Query Match   100.0%; Score 66; DB 16; Length 6166;
Best Local Similarity 100.0%; Pred. No. 6.5e-14;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4541 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 4600

Qy   61 GCGGCC 66
          ||||||
Db   4601 GCGGCC 4606

RESULT 15
AAQ10569
ID   AAQ10569 standard; DNA; 73 BP.
XX
AC   AAQ10569;
XX
DT   26-APR-1991 (first entry)
XX
DE   Omp A signal peptide-encoding sequence.
XX
KW   Mirabilis antiviral protein; MAP; Omp A signal sequence; ds.
XX
OS   Mirabilis jalapa.
XX
PN   EP414134-A.
XX
PD   27-FEB-1991.
XX
PF   16-AUG-1990;   90EP-0115718.
XX
PR   17-AUG-1989;   89JP-0210767.
XX
PA   (NISB) JAPAN TOBACCO INC.
XX
PI   Habuka N, Akiyama K, Tsuge H, Matsumoto T, Noma M;
XX
DR   WPI; 1991-059386/09.
XX
PT   Gene encoding Mirabilis Antiviral Protien and OmpA signal peptide -
PT   for mass-prodn. of extracellular MAP.
XX
PS   Claim 1; page 8; 15pp; English.
XX
CC   This Omp A signal peptide-encoding sequence is contained in a

CC   recombinant plasmid, flanked on its 3' side by a Mirabilis
CC   antiviral protein (MAP) gene. The plasmid is used to transform
CC   E.coli host cells in the large-scale prodn. of MAP. The 5' end
CC   overhangs the 3' end of the complementary strand by TA and the 5'
CC   end of the complementary strand overhangs the 3' end of this sense
CC   strand by CTAG. See also AAQ10568.
XX
SQ   Sequence 73 BP; 16 A; 20 C; 21 G; 16 T; 0 other;

Query Match   91.2%; Score 60.2; DB 12; Length 73;
Best Local Similarity 95.4%; Pred. No. 2.7e-12;
Matches   62; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 61

Qy   61 GCGGC 65
          || ||
Db   62 GCCGC 66

Search completed: August 18, 2003, 23:28:49
Job time : 24.2513 secs

```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 5.16534 Seconds
                                            (without alignments)
                                            5639.757 Million cell updates/sec

Title: US-09-987-455-3
Perfect score: 66
Sequence: 1 atgaaaagacagctatcgc.........ctaccgtggccaggcgcc 66

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
               1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
               2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
               3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
               4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
               5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
               6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	Length	ID	Description
1	66	100.0	3450	4		US-09-318-786-20	Sequence 20, Appl
2	66	100.0	4691	3		US-08-591-632-43	Sequence 43, Appl
3	66	100.0	4691	4		US-09-611-451-43	Sequence 43, Appl
4	66	100.0	6166	3		US-08-591-632-51	Sequence 51, Appl
5	66	100.0	6166	3		US-09-611-451-51	Sequence 51, Appl
6	60.2	91.2	73	1		US-07-854-845B-2	Sequence 2, Appl
7	60.2	91.2	77	3		US-08-666-354A-10	Sequence 10, Appl
8	60.2	91.2	249	1		US-08-439-132-1	Sequence 1, Appl
9	60.2	91.2	249	6		5223407-1	Patent No. 5223407
10	60.2	91.2	548	3		US-08-828-741B-1	Sequence 1, Appl
11	60.2	91.2	548	4		US-09-160-567-1	Sequence 1, Appl
12	60.2	91.2	548	4		US-09-710-299-1	Sequence 1, Appl
13	59.6	90.3	458	2		US-08-880-829-21	Sequence 21, Appl
14	59.6	90.3	470	3		US-08-828-741B-10	Sequence 10, Appl
15	59.6	90.3	470	4		US-09-160-567-10	Sequence 10, Appl
16	59.6	90.3	470	4		US-09-710-299-10	Sequence 10, Appl
17	59.6	90.3	599	3		US-08-828-741B-7	Sequence 7, Appl
18	59.6	90.3	599	4		US-09-160-567-7	Sequence 7, Appl
19	59.6	90.3	599	4		US-09-710-299-7	Sequence 7, Appl
20	59.6	90.3	1031	3		US-08-828-741B-5	Sequence 5, Appl
21	59.6	90.3	1031	4		US-09-160-567-5	Sequence 5, Appl
22	59.6	90.3	1031	4		US-09-710-299-5	Sequence 5, Appl
23	59.6	90.3	1490	3		US-08-828-741B-3	Sequence 3, Appl
24	59.6	90.3	1490	4		US-09-160-567-3	Sequence 3, Appl
25	59.6	90.3	1490	4		US-09-710-299-3	Sequence 3, Appl
26	59.2	89.7	116	1		US-07-842-089B-23	Sequence 23, Appl
27	59.2	89.7	116	1		US-08-264-485-23	Sequence 23, Appl
28	59.2	89.7	245	6		5187153-24	Patent No. 5187153
29	59.2	89.7	245	6		5223482-26	Patent No. 5223482
30	59.2	89.7	1379	4		US-09-553-498-5	Sequence 5, Appli
31	59.2	89.7	1379	4		US-09-618-869-5	Sequence 5, Appli
32	59.2	89.7	1881	4		US-09-553-498-1	Sequence 1, Appli
33	59.2	89.7	1881	4		US-09-553-498-3	Sequence 3, Appli
34	59.2	89.7	1881	4		US-09-618-869-1	Sequence 1, Appli
35	59.2	89.7	1881	2		US-09-618-869-3	Sequence 3, Appli
36	59.2	89.7	4515	1		US-08-256-790-1	Sequence 1, Appli
37	58.8	89.1	63	1		US-07-842-089B-5	Sequence 5, Appli
38	58.8	89.1	63	2		US-08-264-485-5	Sequence 5, Appli
39	58.8	89.1	63	2		US-09-003-081-7	Sequence 7, Appli
40	58.8	89.1	63	3		US-08-648-262-7	Sequence 7, Appli
41	58.8	89.1	63	3		US-08-648-263-7	Sequence 7, Appli
42	58.8	89.1	63	4		US-09-904-196B-4	Sequence 4, Appli
43	58.8	89.1	63	4		US-09-230-233A-6	Sequence 6, Appli
44	58.8	89.1	103	1		US-08-215-138-12	Sequence 12, Appl
45	58.8	89.1	103	1		US-08-407-344-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-318-786-20

```
; Sequence 20, Application US/0931786
; Patent No. 6472147
; GENERAL INFORMATION:
; APPLICANT: Janda, Kim D
; APPLICANT: Wirsching, Peter
; APPLICANT: Lerner, Richard A
; APPLICANT: Gao, Changshou
; TITLE OF INVENTION: METHODS FOR DISPLAY OF HETERODIMERIC PROTEINS ON
; TITLE OF INVENTION: FILAMENTOUS PHAGE USING pVII AND pIX, COMPOSITIONS,
; TITLE OF INVENTION: VECTORS AND COMBINATORIAL LIBRARIES
; FILE REFERENCE: TSR0305S
; CURRENT APPLICATION NUMBER: US/09/318,786
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3450
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Vector
US-09-318-786-20
```

Query Match   100.0%; Score 66; DB 4; Length 3450;
Best Local Similarity   100.0%; Pred. No. 1.1e-13;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ATGAAAAGACAGCTATCGCGATTGCAGTTGCACTGGCACTGGCACTGGCACTGGTTTCGCTACCGTGGCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2667 ATGAAAAGACAGCTATCGCGATTGCAGTTGCACTGGCACTGGCACTGGCACTGGTTTCGCTACCGTGGCCAG 2726

Qy     61 GCGGCC 66
          ||||||
Db   2727 GCGGCC 2732
```

RESULT 2
US-08-591-632-43

```
; Sequence 43, Application US/08591632
; Patent No. 6261558
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F.
; APPLICANT: Burton, Dennis R.
; APPLICANT: Lerner, Righard A.
; TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
```

;   ADDRESSEE: The Scripps Research Institute, Office of
;   ADDRESSEE: Patent Counsel
;   STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8
;   CITY: La Jolla
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92037
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/591,632
;   FILING DATE: 19-OCT-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US94/11907
;   FILING DATE: 19-OCT-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/308,841
;   FILING DATE: 19-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/233,619
;   FILING DATE: 26-APR-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/139,409
;   FILING DATE: 19-OCT-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fitting, Thomas
;   REGISTRATION NUMBER: 34,163
;   REFERENCE/DOCKET NUMBER: TSRI 332.3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 784-2937
;   TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 43:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4691 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
US-08-591-632-43

Query Match   100.0%; Score 66; DB 3; Length 4691;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGCGCACTGGCGGTTTCGCTACCGTGGCCCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGCGCACTGGCGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
            ||||||
Db   2671 GCGGCC 2676

RESULT 3
US-09-611-451-43
; Sequence 43, Application US/09611451
; Patent No. 6395275
; GENERAL INFORMATION:
;   APPLICANT: Barbas, Carlos F.
;   Burton, Dennis R.
;   Lerner, Richard A.
;   TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
;   ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
;   NUMBER OF SEQUENCES: 92
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: The Scripps Research Institute, Office of
;   Patent Counsel
;   STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8
;   CITY: La Jolla
;   STATE: CA

;   COUNTRY: USA
;   ZIP: 92037
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/611,451
;   FILING DATE: 06-Jul-2000
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/591,632
;   FILING DATE: 2001-10-29
;   APPLICATION NUMBER: US 08/308,841
;   FILING DATE: 19-SEP-1994
;   APPLICATION NUMBER: US 08/233,619
;   FILING DATE: 26-APR-1994
;   APPLICATION NUMBER: US 08/139,409
;   FILING DATE: 19-OCT-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fitting, Thomas
;   REGISTRATION NUMBER: 34,163
;   REFERENCE/DOCKET NUMBER: TSRI 332.3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 784-2937
;   TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 43:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4691 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-611-451-43

Query Match   100.0%; Score 66; DB 4; Length 4691;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGCGCACTGGCGGTTTCGCTACCGTGGCCCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 ATGAAAAAGACAGCTATCGCGATTGCAGTGCGCACTGGCGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCGGCC 66
            ||||||
Db   2671 GCGGCC 2676

RESULT 4
US-08-591-632-51
; Sequence 51, Application US/08591632
; Patent No. 6261558
; GENERAL INFORMATION:
;   APPLICANT: Barbas, Carlos F.
;   APPLICANT: Burton, Dennis R.
;   APPLICANT: Lerner, Richard A.
;   TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
;   TITLE OF INVENTION: ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
;   NUMBER OF SEQUENCES: 92
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: The Scripps Research Institute, Office of
;   ADDRESSEE: Patent Counsel
;   STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8
;   CITY: La Jolla
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92037
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/591,632
  FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: PCT/US94/11907
  FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/308,841
  FILING DATE: 19-SEP-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/233,619
  FILING DATE: 26-APR-1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/139,409
  FILING DATE: 19-OCT-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Fitting, Thomas
  REGISTRATION NUMBER: 34,163
  REFERENCE/DOCKET NUMBER: TSRI 332.3
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (619) 784-2937
  TELEFAX: (619) 784-9399
INFORMATION FOR SEQ ID NO: 51:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 6166 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: circular
  MOLECULE TYPE: DNA (genomic)
US-08-591-632-51

Query Match   100.0%;   Score 66;   DB 3;   Length 6166;
Best Local Similarity 100.0%;   Pred. No. 1.2e-13;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4541 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 4600

Qy   61 GCGGCC 66
        ||||||
Db 4601 GCGGCC 4606

RESULT 5
US-09-611-451-51
; Sequence 51, Application US/09611451
; Patent No. 6395275
; GENERAL INFORMATION:
  APPLICANT: Barbas, Carlos F.
           Burton, Dennis R.
           Lerner, Richard A.
  TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
           ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
  NUMBER OF SEQUENCES: 92
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: The Scripps Research Institute, Office of
           Patent Counsel
  STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8
  CITY: La Jolla
  STATE: CA
  COUNTRY: USA
  ZIP: 92037
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/611,451
  FILING DATE: 06-Jul-2000
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/591,632

  FILING DATE: 2001-10-29
  APPLICATION NUMBER: US 08/308,841
  FILING DATE: 19-SEP-1994
  APPLICATION NUMBER: US 08/233,619
  FILING DATE: 26-APR-1994
  APPLICATION NUMBER: US 08/139,409
  FILING DATE: 19-OCT-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Fitting, Thomas
  REGISTRATION NUMBER: 34,163
  REFERENCE/DOCKET NUMBER: TSRI 332.3
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (619) 784-2937
  TELEFAX: (619) 784-9399
INFORMATION FOR SEQ ID NO: 51:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 6166 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: circular
  MOLECULE TYPE: DNA (genomic)
  SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-611-451-51

Query Match   100.0%;   Score 66;   DB 4;   Length 6166;
Best Local Similarity 100.0%;   Pred. No. 1.2e-13;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4541 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 4600

Qy   61 GCGGCC 66
        ||||||
Db 4601 GCGGCC 4606

RESULT 6
US-07-854-845B-2
; Sequence 2, Application US/07854845B
; Patent No. 5340732
; GENERAL INFORMATION:
  APPLICANT: HABUKA, No. 5340732iyuki
  APPLICANT: MIYANO, Masashi
  APPLICANT: MATSUMOTO, Takashi
  APPLICANT: NOMA, Masana
  TITLE OF INVENTION: ANTIVIRAL PROTEIN
  NUMBER OF SEQUENCES: 9
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: BIRCH, STEWART, KOLASCH, & BIRCH
  STREET: 8110 Gatehouse Road, Suite 500 East
  CITY: Falls Church
  STATE: VA
  COUNTRY: USA
  ZIP: 22042
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/854,845B
  FILING DATE: 20-MAR-1992
  CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
  NAME: Murphy Jr., Gerald M
  REGISTRATION NUMBER: 28,977
  REFERENCE/DOCKET NUMBER: 42-194P
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 703-205-8000
  TELEFAX: 703-205-8050
  TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 73 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
;      HYPOTHETICAL: NO
;      ANTI-SENSE: NO
US-07-854-845B-2

Query Match     91.2%;  Score 60.2;  DB 1;  Length 73;
Best Local Similarity  95.4%;  Pred. No. 4.9e-12;
Matches   62;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db    2 ATGAAAAAGACAGCTATCGCGATTGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 61

Qy   61 GCGGC 65
          || ||
Db   62 GCCGC 66


RESULT 7
US-08-666-354A-10
; Sequence 10, Application US/08666354A
; Patent No. 6040141
; GENERAL INFORMATION:
;    APPLICANT: KLAUSER, THOMAS
;    APPLICANT: KRAMER, JOACHIM
;    APPLICANT: MEYER, THOMAS F.
;    APPLICANT: POHLNER, JOHANNES
;    TITLE OF INVENTION: BACTERIA USED TO PRODUCE STABLE FUSION
;    TITLE OF INVENTION: PROTEINS AND METHOD FOR THEIR IDENTIFICATION
;    NUMBER OF SEQUENCES: 18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP.
;      STREET: P.O. BOX 747
;      CITY: FALLS CHURCH
;      STATE: VA
;      COUNTRY: USA
;      ZIP: 22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/666,354A
;      FILING DATE: 23-SEP-1996
;      CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;      NAME: SVENSSON, LEONARD R.
;      REGISTRATION NUMBER: 30,330
;      REFERENCE/DOCKET NUMBER: 147-157P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 703-205-8000
;      TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 77 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: other nucleic acid
;      DESCRIPTION: /desc = "FRAGMENT OF PLASMID pJK165"
US-08-666-354A-10

Query Match     91.2%;  Score 60.2;  DB 3;  Length 77;
Best Local Similarity  95.4%;  Pred. No. 4.9e-12;
Matches   62;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy   61 GCGGC 65
          || ||
Db   61 GCCGC 65


RESULT 8
US-08-439-132-1
; Sequence 1, Application US/08439132
; Patent No. 5646015
; GENERAL INFORMATION:
;    APPLICANT: Wong, W. K. R.
;    APPLICANT: Sutherland, Margaret L.
;    TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS FROM
;    TITLE OF INVENTION: E.COLI
;    NUMBER OF SEQUENCES: 5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Foley & Lardner
;      STREET: Suite 500, 3000 K Street
;      CITY: Washington
;      STATE: D.C.
;      COUNTRY: USA
;      ZIP: 20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/439,132
;      FILING DATE: 11-MAY-1995
;      CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Bent, Stephen A.
;      REGISTRATION NUMBER: 29,768
;      REFERENCE/DOCKET NUMBER: 16777/200/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (202) 672-5300
;      TELEFAX: (202) 672-5399
;      TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 249 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
US-08-439-132-1

Query Match     91.2%;  Score 60.2;  DB 1;  Length 249;
Best Local Similarity  95.4%;  Pred. No. 6.2e-12;
Matches   62;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db  102 ATGAAAAAGACAGCTATCGCGATTGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 161

Qy   61 GCGGC 65
          || ||
Db  162 GCCGC 166


RESULT 9
5223407-1
; Patent No. 5223407
;    APPLICANT: WONG, RAYMOND W.K.;SUTHERLAND, MARGARET L.
;    TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS
; FROM E.COLI
;    NUMBER OF SEQUENCES: 6
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/07/395,797
;      FILING DATE: 18-AUG-1989
```

;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 239,145
;   FILING DATE: 31-AUG-1988
;SEQ ID NO:1:
;   LENGTH: 249
5223407-1

Query Match   91.2%; Score 60.2; DB 6; Length 249;
Best Local Similarity   95.4%; Pred. No. 6.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
            |||||||||||||||||||||||||||||||||||||||||||||| || |||
Db   102 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGGCAG 161

Qy   61 GCGGC 65
            || ||
Db   162 GCCGC 166

RESULT 10
US-08-828-741B-1
; Sequence 1, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
;   APPLICANT: Koentgen, Frank
;   APPLICANT: Suess, Gabriele M.
;   APPLICANT: Tarlinton, David M.
;   APPLICANT: Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
;   TITLE OF INVENTION: PRODUCING SAME
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;   STREET: 400 Garden City Plaza
;   CITY: Garden City
;   STATE: New York
;   COUNTRY: United States of America
;   ZIP: 11530
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/828,741B
;   FILING DATE: 26-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: DiGiglio, Frank S.
;   REGISTRATION NUMBER: 31,346
;   REFERENCE/DOCKET NUMBER: 10591
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 742-4343
;   TELEFAX: (516) 742-4366
;   TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 548 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..548
US-08-828-741B-1

Query Match   91.2%; Score 60.2; DB 3; Length 548;
Best Local Similarity   95.4%; Pred. No. 7.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGGCAG 60
            |||||||||||||||||||||||||||||||||||||||||||||||||| || |||

Qy   61 GCGGC 65
            || ||
Db   61 GCCGC 65

RESULT 11
US-09-160-567-1
; Sequence 1, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
;   APPLICANT: Koentgen, Frank
;   APPLICANT: Suess, Gabriele M.
;   APPLICANT: Tarlinton, David M.
;   APPLICANT: Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
;   TITLE OF INVENTION: PRODUCING SAME
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;   STREET: 400 Garden City Plaza
;   CITY: Garden City
;   STATE: New York
;   COUNTRY: United States of America
;   ZIP: 11530
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/160,567
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/828,741
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: DiGiglio, Frank S.
;   REGISTRATION NUMBER: 31,346
;   REFERENCE/DOCKET NUMBER: 10591
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 742-4343
;   TELEFAX: (516) 742-4366
;   TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 548 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..548
US-09-160-567-1

Query Match   91.2%; Score 60.2; DB 4; Length 548;
Best Local Similarity   95.4%; Pred. No. 7.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
            |||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGGCAG 60

Qy   61 GCGGC 65
            || ||
Db   61 GCCGC 65

RESULT 12

US-09-710-299-1
; Sequence 1, Application US/09710299
; Patent No. 6521741
; GENERAL INFORMATION:
;   APPLICANT: Koentgen, Frank
;   Suess, Gabriele M.
;   Tarlinton, David M.
;   Treutlein, Herbert R.
;   TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
;   PRODUCING SAME
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
;   STREET: 400 Garden City Plaza
;   CITY: Garden City
;   STATE: New York
;   COUNTRY: United States of America
;   ZIP: 11530
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/710,299
;   FILING DATE: 09-No. 6521741-2000
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/828,741
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: DiGiglio, Frank S.
;   REGISTRATION NUMBER: 31,346
;   REFERENCE/DOCKET NUMBER: 10591
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 742-4343
;   TELEFAX: (516) 742-4366
;   TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 548 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..548
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-710-299-1

Query Match   91.2%; Score 60.2; DB 4; Length 548;
Best Local Similarity   95.4%; Pred. No. 7.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCCGGC 65
          || ||
Db   61 GCCGGC 65

RESULT 13
US-08-880-829-21
; Sequence 21, Application US/08880829
; Patent No. 5925559
; GENERAL INFORMATION:
;   APPLICANT: Collins, John
;   APPLICANT: Roettgen, Peter
;   TITLE OF INVENTION: A Collection of Phagemids, A
;   TITLE OF INVENTION: Collection of Escherichia Coli

TITLE OF INVENTION: Cells Carrying The Phagemids, A
TITLE OF INVENTION: Collection of Phagemid Particles
TITLE OF INVENTION: Produced From Said Collection
TITLE OF INVENTION: And Phagemid Particles
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
   ADDRESSEE: Joseph T. Eisele
   ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,
   ADDRESSEE: Levy, Eisele and Richard
   STREET: 711 Third Avenue
   CITY: New York
   STATE: New York
   COUNTRY: U.S.A.
   ZIP: 10017-4059
COMPUTER READABLE FORM:
   MEDIUM TYPE: 3-1/2" DISKETTE
   COMPUTER: IBM-XT COMPATIBLE
   OPERATING SYSTEM: DOS 3.3:
   SOFTWARE: WORDPERFECT 5.0
CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/08/880,829
   FILING DATE: 23-JUN-1997
   CLASSIFICATION: 435
PRIOR APPLICATION DATA:
   APPLICATION NUMBER: 08/458,668
   FILING DATE: 06/02/95
   APPLICATION NUMBER: German EP 94 108 689.4
   FILING DATE: 06/07/94
ATTORNEY/AGENT INFORMATION:
   NAME: EISELE, JOSEPH T.
   REGISTRATION NUMBER: 25,331
   REFERENCE/DOCKET NUMBER: 2727-77
TELECOMMUNICATION INFORMATION:
   TELEPHONE: (212) 687-6000
   TELEFAX: (212) 682-3485
   TELEX: (212) 426767
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
   LENGTH: 458 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single strand
   TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
   ORGANISM:
   STRAIN:
   INDIVIDUAL ISOLATE:
   DEVELOPMENTAL STAGE:
   HAPLOTYPE:
   TISSUE TYPE:
   CELL TYPE:
   CELL LINE:
   ORGANELLE:
IMMEDIATE SOURCE:
US-08-880-829-21

Query Match   90.3%; Score 59.6; DB 2; Length 458;
Best Local Similarity   93.9%; Pred. No. 1.1e-11;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   112 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 171

Qy   61 GCCGGCC 66
          || |||
Db   172 GCCGAC 177

RESULT 14

US-08-828-741B-10
; Sequence 10, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..470
US-08-828-741B-10

Query Match   90.3%;   Score 59.6; DB 3; Length 470;
Best Local Similarity 93.9%;   Pred. No. 1.1e-11;
Matches   62; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db   1 ATGAAAAGACAGCTATCGCGATTGCAGTGCAGTGCACTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy   61 GCGGCC 66
          || ||--|
Db   61 GCCGAC 66

RESULT 15
US-09-160-567-10
; Sequence 10, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..470
US-09-160-567-10

Query Match   90.3%;   Score 59.6; DB 4; Length 470;
Best Local Similarity 93.9%;   Pred No. 1.1e-11;
Matches   62; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db   1 ATGAAAAGACAGCTATCGCGATTGCAGTGCAGTGCACTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy   61 GCGGCC 66
          || ||--|
Db   61 GCCGAC 66

Search completed: August 19, 2003, 08:35:08
Job time : 7.16534 secs

```
              GenCore version 5.1.6

        Copyright (c) 1993 - 2003 Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:    August 19, 2003, 04:46:44 ; Search time 18.6243 Seconds
                               (without alignments)
                        7930.701 Million cell updates/sec

Title:           US-09-987-455-3
Perfect score:   66
Sequence:        1 atgaaaagacagctatgc.........ctaccgtgcccaggcggcc 66

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        1504479 seqs, 118970152 residues

Total number of hits satisfying chosen parameters:    3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries


Database :    Published Applications NA:*
         1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
         2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
         3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
         4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
         5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
         6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
         7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
         8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
         9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
        10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
        11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
        12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
        13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
        14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
        15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
        16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
        17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	
1	66	100.0	66	11	US-09-987-457-1	Sequence 1, Appli
2	66	100.0	66	11	US-09-987-455-3	Sequence 3, Appli
3	66	100.0	66	11	US-09-987-455-6	Sequence 6, Appli
4	66	100.0	1128	11	US-09-987-455-2	Sequence 2, Appli
5	66	100.0	1128	11	US-09-987-455-5	Sequence 5, Appli
6	66	100.0	5149	14	US-10-006-593-60	Sequence 60, Appl
7	66	100.0	6122	14	US-10-006-591-1	Sequence 1, Appli
8	60.2	91.2	548	12	US-10-345-618-1	Sequence 1, Appl
9	59.6	90.3	470	12	US-10-345-618-10	Sequence 10, Appl
10	59.6	90.3	599	12	US-10-345-618-7	Sequence 7, Appl
11	59.6	90.3	932	9	US-09-809-517A-39	Sequence 39, Appl
12	59.6	90.3	1031	12	US-10-345-618-5	Sequence 5, Appli
13	59.6	90.3	1479	12	US-10-345-618-15	Sequence 15, Appl
14	59.6	90.3	1490	12	US-10-345-618-3	Sequence 3, Appli
15	59.6	90.3	1574	9	US-09-809-517A-38	Sequence 38, Appl
16	59.6	90.3	4425	9	US-09-809-517A-40	Sequence 40, Appl
17	59.2	89.7	4145	14	US-10-001-934-36	Sequence 36, Appl
18	59.2	89.7	5020	14	US-10-001-934-35	Sequence 35, Appl
19	59.2	89.7	5079	9	US-09-809-517A-41	Sequence 41, Appl
20	58.8	89.1	63	9	US-09-760-008A-4	Sequence 4, Appli
21	58.8	89.1	63	14	US-10-076-117-3	Sequence 3, Appli
22	58.8	89.1	63	14	US-10-192-294-4	Sequence 4, Appli
23	58.8	89.1	65	9	US-09-875-494-20	Sequence 20, Appl
24	58.8	89.1	108	11	US-09-848-616-8	Sequence 8, Appli
25	58.8	89.1	256	10	US-09-916-230-13	Sequence 13, Appl
26	58.8	89.1	256	11	US-09-848-616-18	Sequence 18, Appl
27	58.8	89.1	261	10	US-09-916-230-15	Sequence 15, Appl
28	58.8	89.1	261	11	US-09-848-616-20	Sequence 20, Appl
29	58.8	89.1	4614	9	US-09-912-165-17	Sequence 17, Appl
30	58.8	87.9	4657	9	US-09-912-165-18	Sequence 18, Appl
31	58	87.9	585	10	US-09-848-585-28	Sequence 28, Appl
32	58	87.9	585	10	US-09-848-585-30	Sequence 30, Appl
33	58	87.9	981	9	US-09-334-477-36	Sequence 36, Appl
34	58	87.9	990	9	US-09-334-477-38	Sequence 38, Appl
35	58	87.9	2321	10	US-09-995-396-2	Sequence 2, Appli
36	58	87.9	2337	10	US-09-995-396-3	Sequence 3, Appli
37	58	87.9	7083	10	US-09-995-396-1	Sequence 1, Appli
38	57.6	87.3	819	12	US-09-226-157-3	Sequence 3, Appl
39	57.2	86.7	903	14	US-10-033-399B-11	Sequence 11, Appl
40 c	55.4	83.9	93	13	US-10-004-832-3	Sequence 3, Appli
41	52.8	80.0	867	11	US-09-782-397-16	Sequence 16, Appl
42 c	52.8	80.0	867	11	US-09-782-397-18	Sequence 18, Appl
43	52.8	80.0	918	11	US-09-782-397-13	Sequence 13, Appl
44 c	52.8	80.0	918	11	US-09-782-397-15	Sequence 15, Appl
45	50	75.8	102	14	US-10-006-593-107	Sequence 107, App

ALIGNMENTS

```
RESULT 1
US-09-987-457-1
; Sequence 1, Application US/09987457
; Publication No. US2003001315OA1
; GENERAL INFORMATION:
;  APPLICANT: Manosroi, Aranya
;  APPLICANT: Manosroi, Jiradej
;  APPLICANT: Tayapiwatana, Chatchai
;  APPLICANT: Goetz, Friedrich
;  APPLICANT: Werner, Rolf-Guenther
;  TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
;  FILE REFERENCE: 0652.2180001
;  CURRENT APPLICATION NUMBER: US/09/987,457
;  CURRENT FILING DATE: 2001-11-14
;  PRIOR APPLICATION NUMBER: 60/268,573
;  PRIOR APPLICATION FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: GB 00 27 782.2
;  PRIOR APPLICATION FILING DATE: 2000-11-14
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;  LENGTH: 66
;  TYPE: DNA
;  ORGANISM: Escherichia coli
US-09-987-457-1

  Query Match             100.0%;  Score 66;  DB 11;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-16;
  Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGTACCGTGGCCCAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66
```

RESULT 2
US-09-987-455-3
; Sequence 3, Application US/0998745S
; Publication No. US2003004972A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 3
;   LENGTH: 66
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-09-987-455-3

Query Match   100.0%; Score 66; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.6e-16;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66

RESULT 3
US-09-987-455-6
; Sequence 6, Application US/0998745S
; Publication No. US2003004972A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 6
;   LENGTH: 66
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-09-987-455-6

Query Match   100.0%; Score 66; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.6e-16;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66

RESULT 4
US-09-987-455-2
; Sequence 2, Application US/0998745S
; Publication No. US2003004972A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
;   LENGTH: 1128
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: coding
;   OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match   100.0%; Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches   66; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCC 66
          ||||||
Db   61 GCGGCC 66

RESULT 5
US-09-987-455-5
; Sequence 5, Application US/0998745S
; Publication No. US2003004972A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
;   LENGTH: 1128

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match   100.0%; Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches   66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy 61 GCCGCC 66
       ||||||
Db 61 GCCGCC 66

RESULT 6
US-10-006-593-60
; Sequence 60, Application US/10006593
; Publication No. US2003004683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 60
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: vector
US-10-006-593-60

Query Match   100.0%; Score 66; DB 14; Length 5149;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches   66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2611 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 2670

Qy   61 GCCGCC 66
        ||||||
Db 2671 GCCGCC 2676

RESULT 7
US-10-006-591-1
; Sequence 1, Application US/10006591
; Publication No. US2003004731A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Renshaw, Mark
; APPLICANT: Wild, Martha
; APPLICANT: McWhirter, John
; TITLE OF INVENTION: ENGINEERED PLASMIDS AND THEIR USE FOR IN SITU PRODUCTION OF GENES

; FILE REFERENCE: 1087-3
; CURRENT APPLICATION NUMBER: US/10/006,591
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/251,440
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
; LENGTH: 6122
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-10-006-591-1

Query Match   100.0%; Score 66; DB 14; Length 6122;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches   66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3050 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 3109

Qy   61 GCCGCC 66
        ||||||
Db 3110 GCCGCC 3115

RESULT 8
US-10-345-618-1
; Sequence 1, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(546)
; OTHER INFORMATION: Description of Artificial Sequence:LHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-1

Query Match   91.2%; Score 60.2; DB 12; Length 548;
Best Local Similarity 95.4%; Pred. No. 1.4e-13;
Matches   62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db   1 ATGAAAAAGACAGCTATCCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy 61 GCCGC 65
       ||| ||
Db 61 GCCGC 65

RESULT 9
US-10-345-618-10
; Sequence 10, Application US/10345618
; Publication No. US20030148484A1

; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(468)
; OTHER INFORMATION: Description of Artificial Sequence:Kappa
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-10

Query Match   90.3%; Score 59.6; DB 12; Length 470;
Best Local Similarity 93.9%; Pred. No. 2.3e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy   61 GCCGCC 66
           |||||
Db   61 GCCGAC 66

RESULT 10
US-10-345-618-7
; Sequence 7, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(597)
; OTHER INFORMATION: Description of Artificial Sequence:LHL.seq
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-7

Query Match   90.3%; Score 59.6; DB 12; Length 599;
Best Local Similarity 93.9%; Pred. No. 2.5e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy   61 GCCGCC 66
           |||||
Db   61 GCCGAC 66

RESULT 11
US-09-809-517A-39
; Sequence 39, Application US/09809517A
; Patent No.US2002003473A1
; GENERAL INFORMATION:
; APPLICANT: Lohning, Corinna
; TITLE OF INVENTION: No. US2002003473Alel methods for displaying (poly)peptides/prote
; TITLE OF INVENTION: particles via disulfide bonds
; FILE REFERENCE: MORPHO/11
; CURRENT APPLICATION NUMBER: US/09/809,517A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: EP 99114072.4
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: EP 00103551.8
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 932
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-39

Query Match   90.3%; Score 59.6; DB 9; Length 932;
Best Local Similarity 93.9%; Pred. No. 2.8e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   266 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 325

Qy   61 GCCGCC 66
           |||||
Db   326 GCCGAC 331

RESULT 12
US-10-345-618-5
; Sequence 5, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; OTHER INFORMATION: Description of Artificial Sequence:TLHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-5

Query Match   90.3%; Score 59.6; DB 12; Length 1031;
Best Local Similarity 93.9%; Pred. No. 2.9e-13;

```
; LOCATION: (1)..(1488)
; OTHER INFORMATION: Description of Artificial Sequence:CATAB-TEV
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-3

Query Match             90.3%;  Score 59.6; DB 12;  Length 1490;
Best Local Similarity   93.9%;  Pred. No. 3.2e-13;
Matches   62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy         61 GCGGCC 66
              || |||
Db         61 GCCGAC 66

RESULT 15
US-09-809-517A-38
; Sequence 38, Application US/09809517A
; Patent No. US2002003473Alel
; GENERAL INFORMATION:
; APPLICANT: Lohning, Corinna
; TITLE OF INVENTION: No. US2002003473Alel methods for displaying (poly)peptides/prote
; TITLE OF INVENTION: particles via disulfide bonds
; FILE.REFERENCE: MORPHO/11
; CURRENT APPLICATION NUMBER: US/09/809,517A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: EP 99114072.4
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: EP 00103551.8
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 1574
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-38

Query Match             90.3%;  Score 59.6; DB 9;  Length 1574;
Best Local Similarity   93.9%;  Pred. No. 3.2e-13;
Matches   62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        266 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 325

Qy         61 GCGGCC 66
              || |||
Db        326 GCCGAC 331

Search completed: August 19, 2003, 14:22:23
Job time : 19.6243 secs
```

```
Matches   62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db          1 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 60

Qy         61 GCGGCC 66
              || |||
Db         61 GCCGAC 66

RESULT 13
US-10-345-618-15
; Sequence 15, Application US/10345618
; Publication No. US2003014848A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:ccMTIgL
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-15

Query Match             90.3%;  Score 59.6; DB 12;  Length 1479;
Best Local Similarity   93.9%;  Pred. No. 3.2e-13;
Matches   62;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||| || |||
Db         28 ATGAAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG 87

Qy         61 GCGGCC 66
              || |||
Db         88 GCCGAC 93

RESULT 14
US-10-345-618-3
; Sequence 3, Application US/10345618
; Publication No. US2003014848A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 22:44:33 ; Search time 193.082 Seconds
(without alignments)
8307.845 Million cell updates/sec

Title: US-09-987-455-3
Perfect score: 66
Sequence: 1 atgaaaagacagctatgc.........ctaccgtgcccaggcggcc 66

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215223056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	48	72.7	655	28	AF075794	AF075794 AF075794
c	2	27.8	42.1	386	14	CB806580	CB806580 AMGNNUC:S
c	3	26.8	40.6	310	9	AI635506	AI635506 ts95a03.x
c	4	26.8	40.6	312	9	AW004057	AW004057 wq85e09.x
c	5	26.8	40.6	321	10	BE503913	BE503913 hz35e09.x
c	6	26.8	40.6	325	9	AA398159	AA398159 zt60e09.s
	7	26.8	40.6	340	9	AA401699	AA401699 zt60e09.r
c	8	26.8	40.6	409	13	BY641012	BY641012 BY641012
	9	26.8	40.6	569	13	BU783334	BU783334 in02g05.y
c	10	26.8	40.6	667	13	BX099539	BX099539 BX099539
c	11	26.8	40.6	694	29	CNS04PHI	CNS04PHI AL301311 Tetraodon
	12	26.8	40.6	786	12	BI560274	BI560274 60353511
c	13	26.8	40.6	871	13	BU853590	BU853590 AGENCOURT
c	14	26.6	40.3	268	28	AQ905715	AQ905715 GSSTc0708
	15	26.6	40.3	359	9	AU244728	AU244728 AU244728
	16	26.6	40.3	377	9	AU244560	AU244560 AU244560
	17	26.6	40.3	450	10	BG234261	BG234261 da39a07.
c	18	26.6	40.3	472	28	AQ904535	AQ904535 GSSTc0415
	19	26.6	40.3	612	12	BJ031326	BJ031326 BJ031326
	20	26.6	40.3	645	13	BQ400448	BQ400448 NISC_mp09
	21	26.6	40.3	753	14	CA808042	CA808042 CA12LI0lI
	22	26.6	40.3	833	14	CB349272	CB349272 CAB2SG000
	23	26.6	40.3	885	13	BQ737041	BQ737041 AGENCOURT
	24	26.6	40.3	888	13	BQ737333	BQ737333 AGENCOURT
	25	26.6	40.3	889	14	CD101043	CD101043 AGENCOURT
	26	26.6	40.3	897	13	BQ733411	BQ733411 AGENCOURT
	27	26.6	40.3	980	14	CB206009	CB206009 AGENCOURT
	28	26.6	40.3	1124	14	CB206058	CB206058 AGENCOURT
	29	26.2	39.7	271	9	AU184837	AU184837 AU184837
c	30	26.2	39.7	374	13	BQ094120	BQ094120 san42d04.
c	31	26.2	39.7	695	14	CD355744	CD355744 UI-M-FY0-
c	32	26.2	39.7	973	13	BQ938965	BQ938965 AGENCOURT
	33	26	39.4	495	28	AQ833969	AQ833969 HS_5261_A
	34	26	39.4	853	29	CNS03TQB	CNS03TQB AL260159 Tetraodon
c	35	26	39.4	950	29	CNS04IET	CNS04IET AL292142 Tetraodon
c	36	26	39.4	1016	29	CNS02VHS	CNS02VHS AL215785 Tetraodon
	37	26	39.4	1387	29	B2557231	B2557231 pacs1-60
	38	25.6	38.8	291	14	H33495	H33495 EST109553_R
	39	25.6	38.8	331	14	CA377775	CA377775 656425 NC
	40	25.6	38.8	438	14	CA373964	CA373964 648233 NC
	41	25.6	38.8	621	9	AL966295	AL966295 AL966295
	42	25.6	38.8	652	9	AL644617	AL644617 AL644617
	43	25.6	38.8	673	9	AL655729	AL655729 AL655729
c	44	25.6	38.8	687	14	CA369617	CA369617 640096 NC
	45	25.6	38.8	691	9	AL630991	AL630991 AL630991

ALIGNMENTS

RESULT 1
AF075794/c
LOCUS   AF075794   655 bp   DNA   linear   GSS 29-AUG-2000
DEFINITION AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella
            typhimurium genomic clone 107-T3, genomic survey sequence.
ACCESSION   AF075794
VERSION   AF075794.1 GI:3320664
KEYWORDS   GSS.
SOURCE   Salmonella typhimurium
ORGANISM   Salmonella typhimurium
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Salmonella.
REFERENCE   1 (bases 1 to 655)
AUTHORS   Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.
TITLE   Sample sequencing of a Salmonella typhimurium LT2 lambda library:
            comparison to the Escherichia coli K12 genome
JOURNAL   FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE   99243757
PUBMED   10227170
COMMENT   Contact: McClelland M
            Molecular Biology
            Sidney Kimmel Cancer Center
            3099 Science Park Road, San Diego, CA 92121, USA
            Email: mcclelland@lifsci.sdsu.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..655

```
                    /organism="Salmonella typhimurium"
                    /mol_type="genomic DNA"
                    /strain="LT2"
                    /db_xref="taxon:602"
                    /clone="107-T3"
                    /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
                    /note="Vector: Lambda DASH II; sequenced using Li-Cor
                    sequencer"
BASE COUNT     195 a     158 c     139 g     161 t     2 others
ORIGIN

Query Match              72.7%;  Score 48;  DB 28;  Length 655;
Best Local Similarity    85.2%;  Pred. No. 8.9e-06;
Matches  52;  Conservative  1;  Mismatches  8;  Indels  0;  Gaps  0;

Qy      5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCAGGCGG 64
          ||||      :||||||||||||||||||||| ||||||||||||||| || |||||| |
Db    298 AAAAAGACNBTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCG 239

Qy     65 C 65
          |
Db    238 C 238

RESULT 2
CB806580/c
LOCUS         CB806580               386 bp    mRNA    linear   EST 16-MAY-2003
DEFINITION    AMGNNUC:SRPG2-00054-G9-A srpg2 (10238) Rattus norvegicus cDNA clone
              srpg2-00054-g9 5', mRNA sequence.
ACCESSION     CB806580
VERSION       CB806580.1  GI:29924779
KEYWORDS      EST.
SOURCE        Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE     1  (bases 1 to 386)
AUTHORS       Amgen EST Program.
TITLE         Amgen Rat EST Program
JOURNAL       Unpublished
COMMENT       Contact: Dan Fitzpatrick
              Amgen, Inc
              One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
              Tel: 805 447-4881
              Plate: 00054 row: g column: 9.
FEATURES      Location/Qualifiers
     source   1..386
              /organism="Rattus norvegicus"
              /mol_type="mRNA"
              /db_xref="taxon:10116"
              /clone="srpg2-00054-g9"
              /tissue_type="peneal gland brain"
              /clone_lib="srpg2 (10238)"
              /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; peneal
              gland brain region"
BASE COUNT      57 a     149 c     119 g     60 t     1 others
ORIGIN

Query Match              42.1%;  Score 27.8;  DB 14;  Length 386;
Best Local Similarity    64.1%;  Pred. No. 48;
Matches  41;  Conservative  0;  Mismatches  23;  Indels  0;  Gaps  0;

Qy      2 TGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGG 61
          ||||| |||||| |||| ||| ||| ||||||||||||||| ||||| |||||||| || ||
Db     85 TGAAGACGGACGGCGCAGCGGCGCGGCGCAGCCGCGCGGCGCTAGGNCAGCGAGCGGCCCCGG 26

Qy     62 CGGC 65
          ||||
Db     25 CGGC 22

RESULT 3
```

```
AI635506/c
LOCUS         AI635506               310 bp    mRNA    linear   EST 14-DEC-1999
DEFINITION    ts95a03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2238988 3',
              mRNA sequence.
ACCESSION     AI635506
VERSION       AI635506.1  GI:4688636
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 310)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL       Unpublished
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
              R. Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
              Bonaldo, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Insert Length: 383    Std Error: 0.00
              Seq primer: -40UP from Gibco.
FEATURES      Location/Qualifiers
     source   1..310
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:2238988"
              /tissue_type="pooled germ cell tumors"
              /lab_host="DH10B"
              /clone_lib="NCI_CGAP_GC6"
              /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
              polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
              from the normalized library NCI CGAP GC4 was prepared, and
              ss circles were made in vitro. Following HAP purification,
              this DNA was used as tracer in a subtractive hybridization
              reaction. The driver was PCR-amplified cDNAs from a pool
              of 5,000 clones made from the same library (cloneIDs
              1257096-1258631, 1469064-1470983, and 1475592-1476743).
              Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      71 a     72 c     64 g     103 t
ORIGIN

Query Match              40.6%;  Score 26.8;  DB 9;  Length 310;
Best Local Similarity    64.5%;  Pred. No. 97;
Matches  40;  Conservative  0;  Mismatches  22;  Indels  0;  Gaps  0;

Qy      5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCCG 64
          ||||    |||||||||||||||||||||| || ||||||||| ||||||||||||||||
Db    268 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTGTTAGCAGAGCCCAGGCAG 209

Qy     65 CC 66
          ||
Db    208 CC 207

RESULT 4
AW004057/c
LOCUS         AW004057               312 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION    wg85e09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2478856 3',
              mRNA sequence.
ACCESSION     AW004057
VERSION       AW004057.1  GI:5850973
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
```

REFERENCE
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 312)
           NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
           R. Emmert-Buck, M.D., Ph.D.
           cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
           Bonaldo, Ph.D.
           cDNA Library Arrayed by: Greg Lennon, Ph.D.
           DNA Sequencing by: Washington University Genome Sequencing Center
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Seq primer: -40UP from Gibco.

FEATURES   Location/Qualifiers
     source   1..312
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2478856"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GC6"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
                     from the normalized library NCI_CGAP_GC4 was prepared, and
                     ss circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     1257096-1258631, 1469064-1470983, and 1475592-1476743).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT   72 a   64 g   72 c   104 t
ORIGIN

Query Match   40.6%; Score 26.8; DB 9; Length 312;
Best Local Similarity   64.5%; Pred. No. 97;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy     5 AAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
           |||| |||| |||||| || |||  || ||||||||| ||  |||| |||||||
Db   269 AAAAAACAGCCATTCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 210

Qy    65 CC 66
           ||
Db   209 CC 208
```

RESULT 5
BE503913/c
LOCUS   BE503913   321 bp   mRNA   linear   EST 04-AUG-2000
            hz35e09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3209992 3',
            mRNA sequence.
ACCESSION   BE503913
VERSION   BE503913.1   GI:9706321
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 321)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 313.

FEATURES   Location/Qualifiers
     source   1..321
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3209992"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GC6"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
                     from the normalized library NCI_CGAP_GC4 was prepared, and
                     ss circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     1257096-1258631, 1469064-1470983, and 1475592-1476743).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT   75 a   77 c   65 g   104 t
ORIGIN

Query Match   40.6%; Score 26.8; DB 10; Length 321;
Best Local Similarity   64.5%; Pred. No. 98;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy     5 AAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
           |||| |||| |||||| || |||  || ||||||||| ||  |||| |||||||
Db   269 AAAAAACAGCCATTCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 210

Qy    65 CC 66
           ||
Db   209 CC 208
```

RESULT 6
AA398159/c
LOCUS   AA398159   325 bp   mRNA   linear   EST 12-AUG-1997
            zt60e09.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726760
            3', mRNA sequence.
ACCESSION   AA398159
VERSION   AA398159.1   GI:2051405
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 325)
AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
TITLE   WashU-Merck EST Project 1997
JOURNAL   Unpublished
COMMENT   Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 392   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham.

FEATURES　　　　Location/Qualifiers
source　　　　　　1..325
　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　/db_xref="GDB:5923670"
　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　/clone="IMAGE:726760"
　　　　　　　　　/sex="male"
　　　　　　　　　/lab_host="DH10B"
　　　　　　　　　/clone_lib="Soares_testis_NHT"
　　　　　　　　　/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
　　　　　　　　　polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
　　　　　　　　　was prepared from mRNA obtained from Clontech Laboratories
　　　　　　　　　, Inc.. and primed with a Not I - oligo(dT) primer [5'
　　　　　　　　　TGTTACCAATCTGAAGTGGGAGCGGCGCGCCCAATTTTTTTTTTTTTTT 3'].
　　　　　　　　　Double-stranded cDNA was ligated to Eco RI adaptors
　　　　　　　　　(Pharmacia), digested with Not I and cloned into the Not I
　　　　　　　　　and Eco RI sites of the modified pT7T3 vector. Library
　　　　　　　　　went through one round of normalization to Cot5, and was
　　　　　　　　　constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT　　　72 a　　　72 c　　　64 g　　　117 t
ORIGIN

Query Match　　　　40.6%;　　　Score 26.8;　DB 9;　Length 325;
Best Local Similarity　64.5%;　　Pred. No. 98;
Matches　40;　Conservative　0;　Mismatches　22;　Indels　0;　Gaps　0;

Qy　　　5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
　　　　　||| |||| |||| ||||| |||| |||||||||| |||| ||||||||||||||||||||
Db　　282 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 223

Qy　　65 CC 66
　　　　　||
Db　222 CC 221

RESULT 7
AA401699
LOCUS　　　　　AA401699　　　　　340 bp　　mRNA　　linear　　EST 12-AUG-1997
DEFINITION　　zt60e09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726760
　　　　　　　　5', mRNA sequence.
ACCESSION　　AA401699
VERSION　　　AA401699.1　GI:2057290
KEYWORDS　　EST.
SOURCE　　　Homo sapiens (human)
ORGANISM　　Homo sapiens
　　　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　　　Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　　1 (bases 1 to 340)
AUTHORS　　Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
　　　　　　　　Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
　　　　　　　　, Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,V., Wylie
　　　　　　　　,T., Waterston,R. and Wilson,R.
TITLE　　　WashU-Merck EST Project 1997
JOURNAL　　Unpublished
COMMENT　　Contact: Wilson RK
　　　　　　　　Washington University School of Medicine
　　　　　　　　4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
　　　　　　　　Tel: 314 286 1800
　　　　　　　　Fax: 314 286 1810
　　　　　　　　Email: est@watson.wustl.edu
　　　　　　　　This clone is available royalty-free through LLNL ; contact the
　　　　　　　　IMAGE Consortium (info@image.llnl.gov) for further information.
　　　　　　　　Insert Length: 392　Std Error: 0.00
　　　　　　　　Seq primer: -28m13 rev2 ET from Amersham.
FEATURES　　　Location/Qualifiers
source　　　　　1..340
　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　/db_xref="GDB:5923670"
　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　/clone="IMAGE:726760"
　　　　　　　　　/sex="male"
　　　　　　　　　/lab_host="DH10B"
　　　　　　　　　/clone_lib="Soares_testis_NHT"
　　　　　　　　　/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
　　　　　　　　　polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
　　　　　　　　　was prepared from mRNA obtained from Clontech Laboratories
　　　　　　　　　Inc.. and primed with a Not I - oligo(dT) primer [5'
　　　　　　　　　TGTTACCAATCTGAAGTGGGAGCGGCGCGCCCAATTTTTTTTTTTTTTT 3'
　　　　　　　　　Double-stranded cDNA was ligated to Eco RI adaptors
　　　　　　　　　(Pharmacia), digested with Not I and cloned into the Not I
　　　　　　　　　and Eco RI sites of the modified pT7T3 vector. Library
　　　　　　　　　went through one round of normalization to Cot5, and was
　　　　　　　　　constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT　　106 a　　　75 c　　　82 g　　　77 t
ORIGIN

Query Match　　　　40.6%;　　　Score 26.8; DB 9;　Length 340;
Best Local Similarity　64.5%;　　Pred. No. 1e+02;
Matches　40;　Conservative　0;　Mismatches　22;　Indels　0;　Gaps　0;

Qy　　　5 AAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
　　　　　||| |||| |||| ||||| |||| |||||||||| |||| ||||||||||||||||||||
Db　　76 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 135

Qy　　65 CC 66
　　　　　||
Db　136 CC 137

RESULT 8
BY641012/c
LOCUS　　　　　BY641012　　　　　409 bp　　mRNA　　linear　　EST 15-DEC-2002
DEFINITION　　BY641012 RIKEN full-length enriched, visual cortex Mus musculus
　　　　　　　　cDNA clone K430341N20 3', mRNA sequence.
ACCESSION　　BY641012
VERSION　　　BY641012.1　GI:26976194
KEYWORDS　　EST.
SOURCE　　　Mus musculus (house mouse)
ORGANISM　　Mus musculus
　　　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　　　Mammalia; Eutheria; Euarchontoglires; Rodentia; Sciurognathi; Muridae; Murinae; Mus
REFERENCE　　1 (bases 1 to 409)
AUTHORS　　Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
　　　　　　　　Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.
　　　　　　　　, Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
　　　　　　　　Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,
　　　　　　　　Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
　　　　　　　　Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani
　　　　　　　　,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
　　　　　　　　,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.
　　　　　　　　, Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
　　　　　　　　Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
　　　　　　　　King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
　　　　　　　　,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
　　　　　　　　,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
　　　　　　　　Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
　　　　　　　　Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
　　　　　　　　,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
　　　　　　　　,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
　　　　　　　　,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,
　　　　　　　　Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa
　　　　　　　　,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
　　　　　　　　Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
　　　　　　　　,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.
　　　　　　　　, Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
　　　　　　　　,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata
　　　　　　　　,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
　　　　　　　　,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE　　Analysis of the mouse transcriptome based on functional annotation
　　　　　　of 60,770 full-length cDNAs
JOURNAL　　Nature 420, 563-573 (2002)
MEDLINE　　22354683
PUBMED　　12466851
COMMENT　　Contact: Yoshihide Hayashizaki

REFERENCE   1 (bases 1 to 569)
AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
             Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
             Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
             Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
             ,M., Gibbons,M., McCann,K., Cole,R., Tsagareishvili,R., Williams,T.
             , Jackson,Y. and Bowers,Y.
TITLE   Endocrine Pancreas Consortium
JOURNAL   Unpublished
COMMENT   Other_ESTs: in02g05.x1
             Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
             Endocrine Pancreas Consortium
             Harvard University, Howard Hughes Medical Institute
             Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
             MA 02138
             Tel: 617-495-1812
             Fax: 617-495-8557
             Email: dmelton@biohp.harvard.edu
             Library was constructed by Dr. J. Ferrer In vivo mass-excised to
             pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
             University Genome Sequencing Center For information on obtaining a
             clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
             Seq primer: -40RP from Gibco
             High quality sequence stop: 429.
FEATURES   Location/Qualifiers
source   1..569
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6123441"
             /tissue_type="insulinoma"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="Human insulinoma"
             /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
             XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
             (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
             pBluescript SK- by Dr. H. Inoue following the Washington
             University protocol
             (http://genome.wustl.edu/est/lambda_protocol.shtml).
             Please contact Hiroshi Inoue, MD/PhD for further
             information on this library (Metabolism Division, Permutt
             Laboratory, Washington University School of Medicine, Box
             8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
             is a Washington University Pancreas EST project library."
BASE COUNT   118 a   160 c   144 g   147 t
ORIGIN

Query Match   40.6%; Score 26.8; DB 13; Length 569;
Best Local Similarity 64.5%; Pred. No. 1.2e+02;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy     5 AAAAGACAGCTATCCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAGGCGG 64
           |||| ||||| ||| || |||| ||| || || |||| ||| || |||||||
Db   253 AAAAACAGCCATCACCAGGCCATTCACGCTTTCACGCTTTGTGCTTTTGTTAGCAGCCCAGGCAG 312

Qy    65 CC 66
           ||
Db   313 CC 314
```

RESULT 10
BX099539
LOCUS   BX099539   667 bp   mRNA   linear   EST 06-FEB-2003
DEFINITION BX099539 Soares_testis_NHT Homo sapiens cDNA clone IMAGp998G231928
            ; IMAGE:782638, mRNA sequence.
ACCESSION   BX099539
VERSION   BX099539.1 GI:27844114
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 667)

            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane
            ,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A.
            , Murata,M., Nakamura,K., Nomura,K., Numazaki,R., Ohno,M., Sakai,K.
            , Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami
            ,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
            Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Tissues were provided by Michela Fagiolini and Takao K. Hensch (
            Laboratory for Neuronal Circuit Development Brain Science Institute
            RIKEN 2-1 Hirosawa,Wako-shi,Saitama 351-0198 Japan) whose
            assistance we gratefully acknowledge.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES   Location/Qualifiers
source   1..409
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6J"
             /db_xref="taxon:10090"
             /clone="K43034N20"
             /tissue_type="visual cortex"
             /clone_lib="RIKEN full-length enriched, visual cortex"
BASE COUNT   89 a   96 c   87 g   135 t   2 others
ORIGIN

Query Match   40.6%; Score 26.8; DB 13; Length 409;
Best Local Similarity 64.5%; Pred. No. 1.1e+02;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy     1 ATGAAAAAGACAGCTATCCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
           | |||||||| ||| || |||| ||| ||| || || |||| ||| || ||||
Db   102 AGGAAAAAGAAGGCAAGCAGGGTTAGAGAGGTGGTGGCTGGGTTCTCAAGCTTTCCACAG 43

Qy    61 GC 62
           ||
Db    42 GC 41
```

RESULT 9
BU783334
LOCUS   BU783334   569 bp   mRNA   linear   EST 11-OCT-2002
DEFINITION in02g05.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123441
            5', mRNA sequence.
ACCESSION   BU783334
VERSION   BU783334.1 GI:23827413
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
           Radelof,U., Schneider,D. and Korn,B.
TITLE   Human UnigeneSet - RZPD3
JOURNAL   Unpublished
COMMENT   Contact: Ina Rolfs
           RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
           Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
           RZPD; IMAGp998G231928.
           RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
           Human Unigeneset - RZPD3 (RZPDLIB No.972)
           http://www.rzpd.de/cloneCards/cgi-
           bin/showLib.pl-cgi/response?libNo=972 Contact: Ina Rolfs
           RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
           Heubnerweg 6, D-14059 Berlin, Germany
           Tel: +49 30 32639 101
           Fax: +49 30 32639 111
           www.rzpd.de
           This clone is available royalty-free from RZPD;
           contact RZPD (clone@rzpd.de) for further information. Seq primer:
           M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES   Location/Qualifiers
     source   1..667
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGp998G231928 ; IMAGE:782638"
                  /sex="male"
                  /lab_host="DH10B"
                  /clone_lib="Soares_testis_NHT"
                  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                  polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
                  was prepared from mRNA obtained from Clontech Laboratories
                  Inc., and primed with a Not I - oligo(dT) primer [5'
                  TGTTACCAATCTGAAGTGGGACGCGGCCGCCCAATTTTTTTTTTTTTT 3'].
                  Double-stranded cDNA was ligated to Eco RI adaptors
                  (Pharmacia), digested with Not I and cloned into the Not I
                  and Eco RI sites of the modified pT7T3 vector. Library
                  went through one round of normalization to Cot5, and was
                  constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 195 a   142 c   167 g   162 t   1 others
ORIGIN

Query Match   40.6%;   Score 26.8;   DB 13;   Length 667;
Best Local Similarity   64.5%;   Pred. No. 1.2e+02;
Matches 40;   Conservative   0;   Mismatches 22;   Indels   0;   Gaps   0;

Qy   5 AAAAGACAGCTATCGGCGATTGCAGTGGCACTGGCACTGGCTTTCGCTACCGTGGCCAGCGG 64
         ||||||||||||||||||| :||||||||||||||||||| | :| || |||||||||||||
Db   43 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 102

Qy   65 CC 66
         ||
Db   103 CC 104

RESULT 11
CNS04PHI/c
LOCUS   CNS04PHI   694 bp   DNA   linear   GSS 01-SEP-2000
DEFINITION   Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
             127D11 of library G from Tetraodon nigroviridis, genomic survey
             sequence.
ACCESSION   AL301311
VERSION   AL301311.1   GI:8178354
KEYWORDS   GSS; genome survey sequence.
SOURCE   Tetraodon nigroviridis
  ORGANISM   Tetraodon nigroviridis
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
             Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,

TITLE   Saurin,W. and Weissenbach,J.
           Estimate of human gene number provided by genome-wide analysis
           using Tetraodon nigroviridis DNA sequence
JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE   20296633
PUBMED   10835645
REFERENCE 2
  AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
           Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
           Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE   Characterization and repeat analysis of the compact genome of the
           freshwater pufferfish Tetraodon nigroviridis
  JOURNAL Genome Res. 10 (7), 939-949 (2000)
  MEDLINE 20359837
  PUBMED   10899143
REFERENCE 3 (bases 1 to 694)
  AUTHORS Genoscope.
  TITLE   Direct Submission
  JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT   This sequence is a single read and was generated as part of a large
           scale clone-end sequencing project of the Tetraodon nigroviridis
           genome. For more information, please take a look at
           http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
     source   1..694
                  /organism="Tetraodon nigroviridis"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:99883"
                  /clone="127D11"
                  /clone_lib="G"
                  /note="Genoscope sequence ID : COBG127CB06SP1-end :
                  PUC-Ori"
BASE COUNT 189 a   122 c   186 g   175 t   22 others
ORIGIN

Query Match   40.6%;   Score 26.8;   DB 29;   Length 694;
Best Local Similarity   60.6%;   Pred. No. 1.3e+02;
Matches 40;   Conservative   2;   Mismatches 24;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCGTTCGCTACCGTGGCCCAG 60
         || ||||||||||||||||| | |||||||| :|| : :|| | || ||||||||||||
Db   278 ACGTKAATGGCGGCTTTTGCGGAARGCTCTKACAATACAKGACGGTATCGTGGCCCAG 219

Qy   61 GCGGCC 66
         ||| ||
Db   218 GCGTCC 213

RESULT 12
BI560274
LOCUS   BI560274   786 bp   mRNA   linear   EST 05-SEP-2001
DEFINITION   603253511F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295746 5',
             mRNA sequence.
ACCESSION   BI560274
VERSION   BI560274.1   GI:15447588
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 786)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
             Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11748 row: f column: 03
High quality sequence stop: 772.

```
FEATURES
     source
          1..786
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:5295746"
          /lab_host="DH10B"
          /clone_lib="NIH_MGC_97"
          /note="Organ: testis; Vector: pBluescriptR (modified
          pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
          ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
          size-selected for  average insert size 2.2 kb and
          normalized to ROT 5. This is a primary library enriched
          for full-length clones and constructed using the
          Cap-trapper method (Carninci, in preparation). Library
          constructed by M. Brownstein (NIMH/NHGRI, National
          Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT     238 a    168 c    187 g    193 t
ORIGIN
Query Match              40.6%;   Score 26.8;  DB 12;  Length 786;
Best Local Similarity    64.5%;   Pred. No. 1.3e+02;
Matches   40;  Conservative   0;  Mismatches  22;  Indels   0;  Gaps   0;

Qy      5 AAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCCGTACCGTGGCCCAGGCGG 64
            ||||| ||||| |||  ||| ||| |||| |||||  ||   |||||||||||||||
Db     56 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGCTTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 115

Qy     65 CC 66
            ||
Db    116 CC 117

RESULT 13
BU853590
LOCUS       BU853590    871 bp    mRNA    linear    EST 16-OCT-2002
DEFINITION  AGENCOURT_10418173 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6620518
            5', mRNA sequence.
ACCESSION   BU853590
VERSION     BU853590.1  GI:24038556
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 871)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2870 row: 1  column: 22
            High quality sequence stop: 454.
FEATURES
     source
          1..871
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:6620518"
          /lab_host="DH10B (T1 phage-resistant)"
          /clone_lib="NIH_MGC_82"
```

```
          /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
          sfiI (ggccgcctggcc); Site_2: SfiI (ggccattatggcc); 5' and
          3' adaptors were used in cloning as follows: 5' adaptor
          sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
          5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
          C, or G and N = A, C, G, or T).  Average insert size
          1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
          inserts by PCR. This library was enriched for full-length
          clones and was constructed by Clontech Laboratories (Palo
          Alto, CA)."
BASE COUNT     240 a    226 c    203 g    202 t
ORIGIN
Query Match              40.6%;   Score 26.8;  DB 13;  Length 871;
Best Local Similarity    64.5%;   Pred. No. 1.4e+02;
Matches   40;  Conservative   0;  Mismatches  22;  Indels   0;  Gaps   0;

Qy      5 AAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCCGTACCGTGGCCCAGGCGG 64
            ||||| ||||| |||  ||| ||| |||| |||||  ||   |||||||||||||||
Db     38 AAAAAACAGCCATCACCAGGCCATTCACGCTTTGCTTTTGTGCTTTTGTTAGCAGAGCCCAGGCAG 97

Qy     65 CC 66
            ||
Db     98 CC 99

RESULT 14
AQ905715/c
LOCUS       AQ905715    268 bp    DNA    linear    GSS 09-JAN-2001
DEFINITION  GSSTc07088 Trypanosoma cruzi random genomic library Trypanosoma
            cruzi genomic clone G55P8, genomic survey sequence.
ACCESSION   AQ905715
VERSION     AQ905715.3  GI:10139418
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi
ORGANISM    Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1  (bases 1 to 268)
AUTHORS     Aguero,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE       A random sequencing approach for the analysis of the trypanosoma
            cruzi genome: general structure, large gene and repetitive DNA
            families, and gene discovery
JOURNAL     Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE     20568489
PUBMED      11116094
COMMENT     On Sep 14, 2000 this sequence version replaced gi:9373324.
            Contact: Sanchez D.O.
            Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
            San Martin)
            Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos
            Aires, Argentina
            Tel: (54-11) 4580/7255/7
            Fax: (54-11) 4752-9639
            Email: dsanchez@iib.unsam.edu.ar
            Sequences were basecalled with phred and vector was masked with
            crossmatch (see http://genome.washington.edu). Sequences were then
            trimmed from both ends to remove low quality bases and masked
            vector.
            Seq primer: T7
            Class: shotgun.
FEATURES
     source
          1..268
          /organism="Trypanosoma cruzi"
          /mol_type="genomic DNA"
          /strain="CL-Brener"
          /db_xref="taxon:5693"
          /clone="G55P8"
          /cell_type="epimastigote"
          /clone_lib="Trypanosoma cruzi random genomic library"
          /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
          randomly sheared using a nebulizer and the 1 to 2 Kb range
          was gel purified and cloned into the dephosphoryated
```

```
BASE COUNT          60 a        63 c        84 g        61 t
ORIGIN                     HincII site of the vector"

Query Match          40.3%;     Score 26.6;  DB 28;  Length 268;
Best Local Similarity  71.4%;   Pred. No. 1.1e+02;
Matches  35;  Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;

Qy    16 ATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCAGGCGG 64
         ||||| ||| ||  ||||| |||| | |  | ||| ||| |||| ||||
Db   189 ATCGTGCTGGCTGTGGCGCGGTGCGCGTTTGGCTCCCGTCGCCCTGGCGG 141

RESULT 15
AU244728
LOCUS       AU244728                  359 bp    mRNA      linear   EST 21-FEB-2002
DEFINITION  AU244728 Shibata Xenopus AEM lambda-ZAP II cDNA library Xenopus
            laevis cDNA clone p5e12 5', mRNA sequence.
ACCESSION   AU244728
VERSION     AU244728.1   GI:18850654
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 359)
  AUTHORS   Shibata,M., Itoh,M., Ohmori,S., Shinga,J. and Taira,M.
  TITLE     Systematic screening and expression analysis of the head organizer
            genes in Xenopus embryos
  JOURNAL   Dev. Biol. 239 (2), 241-256 (2001)
   MEDLINE  21643879
   PUBMED   11784032
COMMENT     Contact: Masanori Taira
            Department of Biological Sciences
            Graduate School of Science, University of Tokyo; CREST, Japan
            Science and Technology Corporation, Japan
            7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
            Tel: 81-03-5841-4434
            Fax: 81-03-5841-4434
            Email: m.taira@biol.s.u-tokyo.ac.jp,
            URL:http://www.biol.s.u-tokyo.ac.jp/users/lmb/lmb-hp.html.
FEATURES             Location/Qualifiers
     source          1..359
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="p5e12"
                     /tissue_type="anterior endomesoderm"
                     /dev_stage="late gastrula to early neurula stages 12.5-13"
                     /clone_lib="Shibata Xenopus AEM lambda-ZAP II cDNA
                     library"
BASE COUNT          90 a        92 c        82 g        93 t       2 others
ORIGIN

Query Match          40.3%;     Score 26.6;  DB 9;  Length 359;
Best Local Similarity  63.1%;   Pred. No. 1.2e+02;
Matches  41;  Conservative  0;  Mismatches  24;  Indels  0;  Gaps  0;

Qy     2 TGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAGG 61
         |||| ||| || ||  ||||| ||| ||  |||||||||| | | | ||||||| ||||
Db   129 TGAAGAACGCAGATATCGGTATCGCTATGGCATCGCTGGTTCTGAATGTCTCCAAGCAGG 188

Qy    62 CGGCC 66
         | |||
Db   189 CAGCC 193

Search completed: August 19, 2003, 08:29:16
Job time : 201.082 secs
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 18, 2003, 19:53:48 ; Search time 4736.31 Seconds
                                          (without alignments)
                                          9743.047 Million cell updates/sec

Title: US-09-987-455-5
Perfect score: 1128
Sequence: 1 atgaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
        1:   gb_ba:*
        2:   gb_htg:*
        3:   gb_in:*
        4:   gb_om:*
        5:   gb_ov:*
        6:   gb_pat:*
        7:   gb_ph:*
        8:   gb_pl:*
        9:   gb_pr:*
       10:   gb_ro:*
       11:   gb_sts:*
       12:   gb_sy:*
       13:   gb_un:*
       14:   gb_vi:*
       15:   em_ba:*
       16:   em_fun:*
       17:   em_hum:*
       18:   em_in:*
       19:   em_mu:*
       20:   em_om:*
       21:   em_or:*
       22:   em_ov:*
       23:   em_pat:*
       24:   em_ph:*
       25:   em_pl:*
       26:   em_ro:*
       27:   em_sts:*
       28:   em_un:*
       29:   em_vi:*
       30:   em_htg_hum:*
       31:   em_htg_inv:*
       32:   em_htg_other:*
       33:   em_htg_mus:*
       34:   em_htg_pln:*
       35:   em_htg_rod:*
       36:   em_htg_mam:*
       37:   em_htg_vrt:*
       38:   em_sy:*
       39:   em_htgo_hum:*
       40:   em_htgo_mus:*
       41:   em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1128	100.0	1128	6	AX431268	AX431268 Sequence
2	1128	100.0	1128	6	AX431271	AX431271 Sequence
3	1065	94.4	1314	6	A27435	A27435 DNA sequenc
4	1065	94.4	1314	6	E01937	E01937 Synthetic D
5	1065	94.4	1341	6	A30593	A30593 DNA for tis
6	1065	94.4	1419	6	A27725	A27725 DNA sequenc
7	1065	94.4	1419	6	AR059996	AR059996 Sequence
8	1065	94.4	1419	6	E01944	E01944 Synthetic D
9	1065	94.4	1689	6	I06609	I06609 Sequence 44
10	1065	94.4	1689	6	I06614	I06614 Sequence 46
11	1065	94.4	1689	6	I08789	I08789 Sequence 3
12	1065	94.4	1780	6	A06611	A06611 Synthetic n
13	1065	94.4	1780	6	A10226	A10226 Synthetic D
14	1065	94.4	1836	6	E01176	E01176 DNA encodin
15	1065	94.4	1836	6	I01581	I01581 Sequence 1
16	1065	94.4	1981	6	I01583	I01583 Sequence 3
17	1065	94.4	1989	6	I01256	I01256 Sequence 2
18	1065	94.4	2099	6	A07207	A07207 Artificial
19	1065	94.4	2100	6	A07197	A07197 H.sapiens m
20	1065	94.4	2100	6	A27429	A27429 cDNA sequen
21	1065	94.4	2100	6	E01934	E01934 DNA encodin
22	1065	94.4	2101	6	AR059986	AR059986 Sequence
23	1065	94.4	2151	6	E00654	E00654 cDNA encodi
24	1065	94.4	2162	6	AR017909	AR017909 Sequence
25	1065	94.4	2162	6	AR017931	AR017931 Sequence
26	1065	94.4	2162	6	I18796	I18796 Sequence 3
27	1065	94.4	2249	6	A07209	A07209 Artificial
28	1065	94.4	2544	6	I08530	I08530 Sequence 1
29	1065	94.4	2547	6	E01163	E01163 cDNA encodi
30	1065	94.4	2547	6	I07841	I07841 Sequence 1
31	1065	94.4	2547	6	I07989	I07989 Sequence 1
32	1065	94.4	2547	6	I09622	I09622 Sequence 1
33	1065	94.4	2549	6	E01055	E01055 DNA encodin
34	1065	94.4	2560	6	A01465	A01465 H.sapiens t
35	1065	94.4	7533	6	I08638	I08638 Sequence 2
36	1064	94.3	1689	6	I08637	I08637 Sequence 1
37	1063.4	94.3	1314	6	AR059989	AR059989 Sequence
38	1063.4	94.3	1506	6	E02247	E02247 cDNA sequen
39 c	1063.4	94.3	1689	6	A14927	A14927 tPA-DNA fro
40	1063.4	94.3	1689	6	A14930	A14930 tPA-DNA fro
41 c	1063.4	94.3	1689	6	A34204	A34204 Human t-PA
42	1063.4	94.3	1689	6	A34205	A34205 Human t-PA
43	1063.4	94.3	1689	6	AX427761	AX427761 Sequence
44	1063.4	94.3	1689	6	E02448	E02448 DNA encodin
45	1063.4	94.3	1689	6	E02663	E02663 DNA encodin

ALIGNMENTS

RESULT 1
AX431268
LOCUS   AX431268   1128 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 2 from Patent WO0240650.
ACCESSION   AX431268
VERSION   AX431268.1   GI:21656150
KEYWORDS   synthetic construct
SOURCE   synthetic construct
ORGANISM   artificial sequences.
REFERENCE   1
AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
TITLE   Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
JOURNAL   Patent: WO 0240650-A 2 23-MAY-2002;

Pred. No. is the number of results predicted by chance to have a

FEATURES
    source   1..1128
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
                    /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT   332 a   332 c   332 g   226 t
ORIGIN

Query Match   100.0%; Score 1128; DB 6; Length 1128;
Best Local Similarity   100.0%; Pred. No. 5.5e-259;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ATGAAAAAGACAGCAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAAAGACAGCAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy    61 GCGGCCTCTGAGGGAAACAGTGACTATCGCGATTGCAGTGCTGCTGCTCTCCAGGCACTG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCGGCCTCTGAGGGAAACAGTGACTATCGCGATTGCAGTGCTGCTGCTCTCCAGGCACTG 120

Qy   121 CACAGCCTCACCGAGTGCGGGTGCCCACGTGCCAAGCCCTGGTTCCACGTGGCCTGATAGGC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACAGCCTCACCGAGTGCGGGTGCCCACGTGCCAAGCCCTGGTTCCACGTGGCCTGATAGGC 180

Qy   181 AAGGTTTACACAGCACAGAACCCAGTGCCACGTGGCCTGGGCACTGGGCAAACATAATTAC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAGGTTTACACAGCACAGAACCCAGTGCCACGTGGCCTGGGCACTGGGCAAACATAATTAC 240

Qy   241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCACGTGCGGCTG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCACGTGCGGCTG 300

Qy   301 ACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAG 360

Qy   361 CCTCAGTTTCGCATCAAAGAGGAGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGCAGGCT 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCTCAGTTTCGCATCAAAGAGGAGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGCAGGCT 420

Qy   421 GCCATCTTTGCCAAGCACAGGAGCGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCATCTTTGCCAAGCACAGGAGCGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTC 480

Qy   481 ATCAGCTCCTGTGGAATTCTCTCGCCCGCGCCCACTGCTTCCGCCCCACGCTGTTCCGCCCCCAC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATCAGCTCCTGTGGAATTCTCTCGCCCGCGCCCACTGCTTCCGCCCCACGCTGTTCCGCCCCCAC 540

Qy   541 CACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGCCCTGGCGAGGAGCAGAAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGCCCTGGCGAGGAGCAGAAA 600

Qy   601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGCACTTACGACAATGAC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGCACTTACGACAATGAC 660

Qy   661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGGATTCGTGCCGCGGACTGGACGCAGCGTGTC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGGATTCGTGCCGCGGACTGGACGCAGCGTGTC 720

Qy   721 CGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCGGACTGGACGGAGTGTGAGCTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCGGACTGGACGGAGTGTGAGCTC 780

Qy   781 TCCGGCTACGGCAAGCATGAGCGCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCT 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TCCGGCTACGGCAAGCATGAGCGCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCT 840

Qy   841 CATGTCAGACTGTACCCATCCAGCCGCTGCACATCAACAACATTTACTTAACAGAACAGTC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CATGTCAGACTGTACCCATCCAGCCGCTGCACATCAACAACATTTACTTAACAGAACAGTC 900
```

```
Qy   901 ACCGACAACATGTCGTGTGCTGGAGACACTCGGAGCGGCGCGGGCCCCCAGGCAAACTTGCAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 ACCGACAACATGTCGTGTGCTGGAGACACTCGGAGCGGCGCGGGCCCCCAGGCAAACTTGCAC 960

Qy   961 GACGGCCTGCCAGGGCGATTCGGGAGGCCCTGGGCGTGGGCCTGGGCGTGTCTGAACGATGGCCGCATGACT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GACGGCCTGCCAGGGCGATTCGGGAGGCCCTGGGCGTGGGCCTGGGCGTGTCTGAACGATGGCCGCATGACT 1020

Qy  1021 TTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTTGGACAGAAGGAGTGTCCGGGTGTGTAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTTGGACAGAAGGAGTGTCCGGGTGTGTAC 1080

Qy  1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

RESULT 2
AX431271
LOCUS   AX431271   1128 bp   DNA   linear   PAT 28-JUN-2002
DEFINITION Sequence 5 from Patent WO0240650.
ACCESSION   AX431271
VERSION   AX431271.1 GI:21656153
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE   Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 5 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES   Location/Qualifiers
     source   1..1128
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT   238 a   332 c   332 g   226 t
ORIGIN

Query Match   100.0%; Score 1128; DB 6; Length 1128;
Best Local Similarity   100.0%; Pred. No. 5.5e-259;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ATGAAAAAGACAGCAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAAAGACAGCAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy    61 GCGGCCTCTGAGGGAAACAGTGACTATCGCGATTGCAGTGCTGCTGCTCTCCAGGCACTG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCGGCCTCTGAGGGAAACAGTGACTATCGCGATTGCAGTGCTGCTGCTCTCCAGGCACTG 120

Qy   121 CACAGCCTCACCGAGTGCGGGTGCCCACGTGCCAAGCCCTGGTTCCACGTGGCCTGATAGGC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACAGCCTCACCGAGTGCGGGTGCCCACGTGCCAAGCCCTGGTTCCACGTGGCCTGATAGGC 180

Qy   181 AAGGTTTACACAGCACAGAACCCAGTGCCACGTGGCCTGGGCACTGGGCAAACATAATTAC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAGGTTTACACAGCACAGAACCCAGTGCCACGTGGCCTGGGCACTGGGCAAACATAATTAC 240

Qy   241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCACGTGCGGCTG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCACGTGCGGCTG 300

Qy   301 ACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACAGCCAG 360

Qy   361 CCTCAGTTTCGCATCAAAGAGGAGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGCAGGCT 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCTCAGTTTCGCATCAAAGAGGAGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGCAGGCT 420
```

```
Db    361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCT 420

ORIGIN

      Query Match           94.4%;    Score 1065;  DB 6;    Length 1314;
      Best Local Similarity 99.5%;    Pred. No. 6.5e-244;
      Matches 1068;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     56 CCCAGGCGGCCTCTGAGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTG 115
           |||.||.||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTG 298

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGAGGAGGAGCAGAAA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGAGGAGGAGCAGAAA 358

Qy    176 TAGGCAAGGTTTACACAGGACAGACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 TAGGCAAGGTTTACACAGGACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 418

Qy    236 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 478

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCCCTGCTCAAGGACCAGTACA 355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCCCTGCTCAAGGACCAGTACA 538

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 598

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGACAGAGGTGCCCGGACAGTC 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    599 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGACAGAGGTGCCCGGACAGTC 658

Qy    476 TACTCATCAGCTCCTGCTCTCTGCGATTCTCTCTGCGGGTTTCCGC 535
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    659 TACTCATCAGCTCCTGCTCTCTGCGATTCTCTCTGCGGGTTTCCGC 718

Qy    536 CCCCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGAGGAGGAGC 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    719 CCCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCGAGGAGGAGC 778

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    779 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 838

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGCAGCG 898

Qy    716 TGGTCCGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGAGCGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 TGGTCCGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGAGCGGAGTGTG 958

Qy    776 AGCTTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 AGCTTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGG 1018

Qy    836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACATTACTTAACGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACATTACTTAACGAA 1078

Qy    896 CAGTACCGACAACATGCTGTGTGCTGAGACACTCCGAGAGCGGCGGCCCCAGGCAAACT 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 CAGTACCGACAACATGCTGTGTGCTGAGACACTCCGAGAGCGGCGGCCCCAGGCAAACT 1138

Qy    956 TGCACGACGCCTGCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1139 TGCACGACGCCTGCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1198

Qy   1016 TGACTTTTGGTGGGCATCATCAGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1199 TGACTTTTGGTGGGCATCATCAGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1258
```

(column 2)

```
Db    421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGACAGAGCGGTTCCGTGCGGGGGCATACTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGACAGAGCGGTTCCGTGCGGGGGCATACTC 480

Qy    481 ATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCGAGGAGGAGGTTTCCGCCCCAC 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCGAGGAGGAGGTTTCCGCCCCAC 540

Qy    541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAA 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGCAGAAA 600

Qy    601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC 660

Qy    661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGCAGCGTGTC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGTGCCCAGGAGCAGCGTGTC 720

Qy    721 CGGCTACGGCCAAGCATGAGGCCTTCCCCGGGGACCTGCAGCTGCCGGACTGGAGCTC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CGGCTACGGCCAAGCATGAGGCCTTCCCCGGGGACCTGCAGCTGCCGGACTGGAGCTC 780

Qy    781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCT 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGGAGGCT 840

Qy    841 CATGTCAGACTGTACCATCCAGCCGCTGCACATCACACATTACTTAACGAACACAGTC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATGTCAGACTGTACCATCCAGCCGCTGCACATCACACATTACTTAACGAACACAGTC 900

Qy    901 ACCGACAACATGCTGTGTGCTGAGACACTCCGAGAGCGGCGGCCCCCAGGCAAACTTGCAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ACCGACAACATGCTGTGTGCTGAGACACTCCGAGAGCGGCGGCCCCCAGGCAAACTTGCAC 960

Qy    961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACT 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACT 1020

Qy   1021 TTGGTGGGCATCATCAGCTGGGGCTGGGCCTGTGGACAGAAGGATGTCCCGGGTGTGTAC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTGGTGGGCATCATCAGCTGGGGCTGGGCCTGTGGACAGAAGGATGTCCCGGGTGTGTAC 1080

Qy   1081 ACAAAGGTTACCAACTACCTAGACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACAAAGGTTACCAACTACCTAGACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

```
RESULT 3
A27435
LOCUS       A27435              1314 bp    DNA     linear        PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pTQkPAdeltatrp.
ACCESSION   A27435
VERSION     A27435.1   GI:21727230
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 032456-A 35 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1314
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      286 a    386 c    393 g    249 t
```

```
Qy    1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1259  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1311
```

RESULT 4
```
LOCUS       E01937
DEFINITION  Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01937
VERSION     E01937.1  GI:2170185
KEYWORDS    JP 1989104167-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 4 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT  OS   Artificial gene
         OC   Artificial sequence; Genes.
         PN   JP 1989104167-A/4
         PD   21-APR-1989
         PF   01-AUG-1988 JP 1988192320
         PR   03-AUG-1987 GB 87  8718298,  26-OCT-1987 GB 87  8725052, PR
         13-NOV-1987 GB 87  8726683
         PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
         PI   NOTANI JOJI,
         PI   KOBAYASHI MASAKAZU
         PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00,C12N15/00//A61K37/54,(C12N9/64, PC
         C12R1:19),
         PC   (C12N9/64,C12R1:91);
         CC   strandedness: Double;
         CC   topology: Linear;
         CC   hypothetical: No;
         CC   anti-sense: No;
         CC   *source: clone=pTQkPAdeltatrp;
         FH   Key             Location/Qualifiers
         FH
         FT   CDS             1..1314
         FT                   /product='t-PA'.
FEATURES             Location/Qualifiers
     source          1..1314
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT    286 a    386 c    393 g    249 t
ORIGIN
```

```
Query Match  94.4%;   Score 1065;   DB 6;   Length 1314;
Best Local Similarity 99.5%;   Pred. No. 6.5e-244;
Matches 1068;   Conservative 0;   Mismatches 5;   Indels 0;   Gaps 0;
```

```
Qy    56   CCCAGGCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239   CCCCTGCCTGCTCTGAGGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  298

Qy   116   GCACGCACCAGCCTCACCGAGTCGGGTGCCTCCGGTGCCTCCCGTGGGACTGGATCCTGA  175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299   GCACGCACCAGCCTCACCGAGTCGGGTGCCTCCGGTGCCTCCCGTGGGACTGGATCCTGA  358

Qy   176   TAGGCAAGGTTTACACAGCACAGAGAACCCCAGTGCCCAGGGCACTGGGCCTGGGCAAACATA  235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359   TAGGCAAGGTTTACACAGCACAGAGAACCCCAGTGCCCAGGGCACTGGGCCTGGGCAAACATA  418

Qy   236   ATTACTGCCCGGAATCCTGATGGGGATGCCAAGCCCTGCTGCCACGTGCCACGTGCCTGAAGAACCGCA  295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419   ATTACTGCCCGGAATCCTGATGGGGATGCCAAGCCCTGCTGCCACGTGCCACGTGCCTGAAGAACCGCA  478

Qy   296   GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  355
```

```
Qy   479  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356  GCCAGCCTCAGTTTCGCATCAAAGAGCGGCTCTTCGCCGCACATCGCCTCCTGGC  415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539  GCCAGCCTCAGTTTCGCATCAAAGAGCGGCTCTTCGCCGCACATCGCCTCCTGGC  598

Qy   416  AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGGCCGGAGAGGCGGTTCCTGTGCGGGGCA  475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599  AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGGCCGGAGAGGCGGTTCCTGTGCGGGGCA  658

Qy   476  TACTCATCAGCTCCTGGATTCTCTCTGCGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGAGC  535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659  TACTCATCAGCTCCTGGATTCTCTCTGCGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGAGC  718

Qy   536  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGAGC  595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   719  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGAGC  778

Qy   596  AGAAATTTGAAGTCGAAAAATCATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   779  AGAAATTTGAAGTCGAAAAATCATTGTCCATAAGGAATTCGATGATGACACTTACGACA  838

Qy   656  ATGACATTGCGCTGCTCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCCG  715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   839  ATGACATTGCGCTGCTCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCCG  898

Qy   716  TGGTCCGACACTGTGTGCCTTCCCCGGCGAGCTGACCTGCCAGCTGACGCGGAGTGTG  775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   899  TGGTCCGACACTGTGTGCCTTCCCCGGCGAGCTGACCTGCCAGCTGACGCGGAGTGTG  958

Qy   776  AGCTCTCCGGCTACGGCAAGCAAGCATGAGGCCTTGTGCTCCTTTCATTCGGAGCGGCTGAAGG  835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   959  AGCTCTCCGGCTACGGCAAGCAAGCATGAGGCCTTGTGCTCCTTTCATTCGGAGCGGCTGAAGG  1018

Qy   836  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACCAGAA  895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACCAGAA  1078

Qy   896  CAGTCACCGACAACATGTCTGTGTCTGGAGACACTCGGAGACACTCGGAGCGGCGGGCCCAGGCAAACT  955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 CAGTCACCGACAACATGTCTGTGTCTGGAGACACTCGGAGACACTCGGAGCGGCGGGCCCAGGCAAACT  1138

Qy   956  TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGTGTGTCTGAACGATGGCCGCA  1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1139 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGTGTGTCTGAACGATGGCCGCA  1198

Qy   1016 TGACTTTGGTGGCATCATCAGCTGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG  1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1199 TGACTTTGGTGGCATCATCAGCTGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG  1258

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1259 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1311
```

RESULT 5
```
LOCUS       A30593          1341 bp   DNA   linear   PAT  04-AUG-1995
DEFINITION  DNA for tissue plasminogen activator (t-PA) from patent EP0400545.
ACCESSION   A30593
VERSION     A30593.1  GI:1249165
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1341)
  AUTHORS   Stern,A., Kohnert,U., Rudolph,R., Fischer,S. and Martin,U.
  TITLE     Derivative of tissue plasminogen activator
  JOURNAL   Patent: EP 0400545-A 2 05-DEC-1990;
            BOEHRINGER MANNHEIM GMBH
FEATURES             Location/Qualifiers
     source          1..1341
```

```
RESULT 6
A27725
LOCUS       A27725                    1419 bp    DNA        linear   PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pmTQk112.
ACCESSION   A27725
VERSION     A27725.1   GI:2172736
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 49 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1419
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      309 a    413 c    426 g    271 t
ORIGIN
```

Query Match   94.4%; Score 1065; DB 6; Length 1419;
Best Local Similarity   99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

```
Qy   56 CCCAGGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 CCCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 403

Qy  116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  404 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 463

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  464 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 523

Qy  236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  524 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCGAAGAACCGCA 583

Qy  296 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  584 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA 643

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  644 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 703

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGGAGTCGCGGGGGGA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  704 AGGCTGCCATCTTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGGAGTCGCGGGGGGCA 763

Qy  476 TACTCATCAAGTCCTGGATTCTCTGCCTGAGATTCTCTGCCTGAGATTCCAGGAGAGGTTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  764 TACTCATCAAGTCCTGGATTCTCTGCCTGAGATTCTCTGCCTGAGATTCCAGGAGAGGTTTCCGC 823

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGTTCCCTGGCGGGGGAGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  824 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGTTCCCTGGCGGGGGAGC 883
```

/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
```
BASE COUNT      293 a    397 c    399 g    252 t
ORIGIN
```

Query Match   94.4%; Score 1065; DB 6; Length 1341;
Best Local Similarity   99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

```
Qy   56 CCCAGGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  269 CCCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 328

Qy  116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  329 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 388

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  389 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 448

Qy  236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  449 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCGAAGAACCGCA 508

Qy  296 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  509 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA 568

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  569 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCCTCCCACCCCTGGC 628

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGGAGTCGCGGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  629 AGGCTGCCATCTTTGCCAAGCACAGGAGGAGTCGCCCGGAGAGGAGTCGCGGGGGGCA 688

Qy  476 TACTCATCAGTCCTGCTGATTCTCTGCGCCGGGTGTCGCCCGCCCCACCGGGTTTCCGC 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  689 TACTCATCAGTCCTGCTGATTCTCTGCGCCGGGTGTCGCCCGCCCCACCGGGTTTCCGC 748

Qy  536 CCCACCACCTGACGGTGATGTCTGGGCAGAAACATACCGGGTCCCTGGCGAGGAGGAGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  749 CCCACCACCTGACGGTGATGTCTGGGCAGAAACATACCGGGTCCCTGGCGAGGAGGAGAGC 808

Qy  596 AGAAATTTGAAGTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  809 AGAAATTTGAAGTCGAAAAATACATTGTCCATTAAGGAATTCGATGATGACACTTACGACA 868

Qy  656 ATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGGCGTGCCCAGGAGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  869 ATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGGCGTGCCCAGGAGAGCAGCG 928

Qy  716 TGGTCCGCACTGTGTGCCTTCCCCCGGCAGCCTGACTGGACGGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  929 TGGTCCGCACTGTGTGCCTTCCCCCGGCAGCCTGACTGGACGGAGTGTG 988

Qy  776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCCGGAGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  989 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCCGGAGGCTGAAGG 1048

Qy  836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1049 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1108

Qy  896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCCGCGGCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1109 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCCGCGGCCCCAGGCAAACT 1168

Qy  956 TGCACGACGCCTCTGCCAGGGCGATTCGGGAGGCCCCCTTGTGTCTGAACGATGGCCGCA 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1169 TGCACGACGCCTCTGCCAGGGCGATTCGGGAGGCCCCCTTGTGTCTGAACGATGGCCGCA 1228

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGGCCTGGGCGGTGTGGACAGAAGAAGGATGTCCCGGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1229 TGACTTTGGTGGGCATCATCAGCTGGGGGCCTGGGCGGTGTGGACAGAAGAAGGATGTCCCGGGGTG 1288

Qy  1076 TGTACACAAAGGTTACCACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1289 TGTACACAAAGGTTACCACTACCTAGACTGGATTCGTGACAACATGCGACCG 1341
```

```
Qy    596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy    656 ATGACATTGCGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCGG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 ATGACATTGCGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCGG 1003

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGGACGGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGGACGGAGTGTG 1063

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1123

Qy    836 AGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 AGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1183

Qy    896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCCCAGGCAAACT 1243

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1303

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 TGACTTTGGTGGGCATCATCAGCTGGGGGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1363

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1416
```

```
RESULT 7
AR059996
LOCUS       AR059996                1419 bp    DNA       linear    PAT 29-SEP-1999
DEFINITION  Sequence 62 from patent US 5840533.
ACCESSION   AR059996
VERSION     AR059996.1  GI:5986446
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 62 24-NOV-1998;
FEATURES    Location/Qualifiers
     source          1..1419
                     /organism="unknown"
BASE COUNT      309 a    416 c    423 g    271 t
ORIGIN
```

```
Query Match     94.4%;  Score 1065;  DB 6;  Length 1419;
Best Local Similarity   99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy     56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
          |||  ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 CCCCTGCCTCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 403

Qy    116 GCACGCACAGCCTCACCGAGCTCGGGTGCCTCCTGCGTCCTCCGTGGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 GCACGCACAGCCTCACCGAGCTCGGGTGCCTCCTGCGTCCTCCGTGGAATTCCATGATCCTGA 463

Qy    176 TAGGCAAGGTTTACAGCACAGCACAGAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    464 TAGGCAAGGTTTACAGCACAGCACAGAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 523
Qy    236 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 583

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGTCGTCCACCTGCGGCCTGAGACAGTACA 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    584 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGTCGTCCACCTGCGGCCTGAGACAGTACA 643

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    644 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 703

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGGCGGTTCGTGTGGGGGGCCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGGCGGTTCGTGTGGGGGGCCA 763

Qy    476 TACTCATCAGCTCCTGCGATCTTCTCTGCCACTGCTTCCAGGAGAGGTTTCCGC 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 TACTCATCAGCTCCTGCGATCTTCTCTGCCACTGCTTCCAGGAGAGGTTTCCGC 823

Qy    536 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGCCTGGTGGGAGGAGGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    824 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCCGGAGGAGGAGC 883

Qy    596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy    656 ATGACATTGCGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCGG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 ATGACATTGCGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCGG 1003

Qy    716 TGGTCCGGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGGTCCGGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGGACGGAGTGTG 1063

Qy    776 AGCTCTCCGGGTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 AGCTCTCCGGGTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1123

Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1183

Qy    896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCCCAGGCAAACT 1243

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1303

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 TGACTTTGGTGGGCATCATCAGCTGGGGGCTGGGGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1363

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1416
```

```
RESULT 8
E01944
LOCUS       E01944                  1419 bp    DNA       linear    PAT 29-SEP-1997
DEFINITION  Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01944
VERSION     E01944.1  GI:2170192
KEYWORDS    JP 198910192  A/11.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
```

```
REFERENCE    1 (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 11 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
  COMMENT   OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/11
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19),
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pmTQk112;
            FH   Key           Location/Qualifiers
            FH
            FT   CDS           1..1419
            FT                 /product='new t-PA'.
FEATURES             Location/Qualifiers
     source          1..1419
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      309 a    413 c    426 g    271 t
ORIGIN

Query Match             94.4%;  Score 1065;  DB 6;  Length 1419;
Best Local Similarity   99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy     56 CCCAGGCGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 403

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGAATTCCATGATCCTGA 463

Qy    176 TAGGCAAGGTTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 TAGGCAAGGTTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 523

Qy    236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCCA 583

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    584 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 643

Qy    356 GCCAGCCTCAGTTTCGCATCAAAAGGAGGGCTCTTCGCCGACATCGCCTTCCCACCCCTGGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    644 GCCAGCCCTCAGTTTCGCATCAAAAGGAGGGCTCTTCGCCGACATCGCCTTCCCACCCCTGGC 703

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTCGCGGGGGCCA 475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTCGCGGGGGCA 763

Qy    476 TACTCATCAGCTCCCTGCTGGAATTCTCTCTGCCGCCCCACTGCTTCCAGGAGAGTTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 TACTCATCAGCTCCCTGCTGGAATTCTCTCTGCCGCCCCACTGCTTCCAGGAGAGTTTCCGC 823

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCCGGGTGGTCCCCTGCGGGAGGAGC 595
```

```
Db    824 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCCGGGTGGTCCCCTGCGGGAGGAGC 883

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1003

Qy    716 TGGTCCGCACTGTGTCCTTCCCCGGGACCTTCAGCTGCCGGACCTGACGGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGGTCCGCACTGTGTCCTTCCCCGGGACCTTCAGCTGCCGGACCTGACGGAGTGTG 1063

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTCATTCGGAGCGGCTGAAGG 1123

Qy    836 AGGCTCATGTGACACTGTACCCATCCAGCCGCGCTGCACATCACACATTACTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 AGGCTCATGTGACACTGTACCCATCCAGCCGCGCTGCACATCACACATTACTAACAGAA 1183

Qy    896 CAGTCACCGACACACATGCTGTGTCTGAGACACTCGGAGCGGGGCCCCAGCCAAACT 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 CAGTCACCGACACACATGCTGTGTCTGAGACACTCGGAGCGGGGCCCCAGCCAAACT 1243

Qy    956 TGCACGACGCCTGCCAGGCGGATTCGGGAGGCCCCTGGTGTGTTCTGAACGATGGCCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 TGCACGACGCCTGCCAGGCGGATTCGGGAGGCCCCTGGTGTGTTCTGAACGATGGCCCGCA 1303

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTGTGACAGAAGGATGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 TGACTTTGGTGGGCATCATCAGCTGGGGCTGTGACAGAAGGATGTCCCGGGTG 1363

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1416
```

```
RESULT 9
I06609
I06609
LOCUS       I06609          1689 bp    DNA    linear   PAT 02-DEC-1994
DEFINITION  Sequence 44 from Patent EP 0293934.
ACCESSION   I06609
VERSION     I06609.1  GI:590639
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1689)
  AUTHORS   Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
            Hashimoto,A. and Yuzuriha,T.
  TITLE     Mutant t-PA with kringle replacement
  JOURNAL   Patent: EP 0293934-A1 44 07-DEC-1988;
FEATURES            Location/Qualifiers
     source         1..1689
                    /organism="unknown"
BASE COUNT     377 a    483 c    505 g    324 t
ORIGIN

Query Match             94.4%;  Score 1065;  DB 6;  Length 1689;
Best Local Similarity   99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy     56 CCCAGGCGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     56 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 673

Qy    116 GCACGCAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614 GCACGCAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGAATTCCATGATCCTGA 673

Qy    116 GCACGCACAGCAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTCCTGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    674 GCACGCACAGCAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTCCTGGAATTCCATGATCCTGA 733
```

ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1689)
AUTHORS   Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
            Hashimoto,A. and Yuzuriha,T.
TITLE   Mutant t-PA with kringle replacement
JOURNAL   Patent: EP 0293934-A1 46 07-DEC-1988;
FEATURES   Location/Qualifiers
   source   1..1689
            /organism="unknown"
BASE COUNT 377 a   484 c   504 g   324 t
ORIGIN

Query Match   94.4%;   Score 1065; DB 6;   Length 1689;
Best Local Similarity   99.5%;   Pred. No. 6.5e-244;
Matches 1068; Conservative   0; Mismatches   0; Indels   5; Gaps   0;

```
Qy      56 CCCAGGCGGCCTCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     614 CCCCTGCCTGCTGCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 673

Qy     116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     674 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA 733

Qy     176 TAGGCAAGGTTTACACAGACACAGAACCCCAGTGCCCAGTGCCTGGGCACTGGGCGAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     734 TAGGCAAGGTTTACACAGACACAGAACCCCAGTGCCCAGTGCCTGGGCACTGGGCAAACATA 793

Qy     236 ATTACTGCCGGAATCCTGATGGGATGGGACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     794 ATTACTGCCGGAATCCTGATGGGATGGGACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 853

Qy     296 GGCTGACGTGGGAGTACTGTGATCGGGTGGTCCCTGCGTGGGAGGAGC 355
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     854 GGCTGACGTGGGAGTACTGTGATCGGGTGGTCCCTGCGTGGGAGGAGC 913

Qy     356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     914 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 973

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     974 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGGCA 1033

Qy     476 TACTCATCAGCTCCTGCTTCTCTGCGCCCACTGCTTCTCCAGGAGAGGTTTCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1034 TACTCATCAGCTCCTGCTTCTCTGCGCCCACTGCTTCTCCAGGAGAGGTTTCCGC 1093

Qy     536 CCCACCACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGCCCTGCGTGGTCCCTGCGGAGGAGC 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1094 CCCACCACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGCCCTGCGTGGTCCCTGCGGAGGAGC 1153

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACA 1213

Qy     656 ATGACATTGCGCTGCGCAGGTAGTTCGTGCGCCGGAACACTCGGAGACACTCGGGTGCGTGTCGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1214 ATGACATTGCGCTGCGCAGGTAGTTCGTGCGCCGGAACACTCGGAGACACTCGGGTGCGTGTCGCAGCG 1273

Qy     716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGAGACCTGCAGCTGACCGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1274 TGGTCCGCACTGTGTGCCTTCCCCCGGCGAGACCTGCAGCTGACCGGAGTGTG 1333

Qy     776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1393

Qy     836 AGCTCATGTCAGACTGTACCCATCCAGCCGCTGCAACATCACAACATTTACTTAACAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1394 AGCTCATGTCAGACTGTACCCATCCAGCCGCTGCAACATCACAACATTTACTTAACAGAA 1453
```

RESULT 10
I06614
LOCUS   I06614   1689 bp   DNA   linear   PAT 02-DEC-1994
DEFINITION Sequence 46 from Patent EP 0293934.
ACCESSION   I06614
VERSION   I06614.1 GI:590640
KEYWORDS   .
SOURCE   Unknown.

(Left column alignment — rotated / second panel:)

```
Qy     176 TAGGCAAGGTTTACACAGACACAGAACCCCAGTGCCCAGTGCCTGGGCACTGGGCGAAACATA 235
Db     734 TAGGCAAGGTTTACACAGACACAGAACCCCAGTGCCCAGTGCCTGGGCACTGGGCAAACATA 793

Qy     236 ATTACTGCCGGAATCCTGATGGGATGGGACAGTGCCGTGCCCAGTGCCTGGGCGAAGAACCGCA 295
Db     794 ATTACTGCCGGAATCCTGATGGGATGGGACAGTGCCGTGCCCAGTGCCTGGGCGAAGAACCGCA 853

Qy     296 GGCTGACGTGGGAGTACTGTGATCGGGTGGTCCCTGCGTGGGAGGAGC 355
Db     854 GGCTGACGTGGGAGTACTGTGATCGGGTGGTCCCTGCGTGGGAGGAGC 913

Qy     356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGC 415
Db     914 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGC 973

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGCGGTTCCTGTGCGGGGGCA 475
Db     974 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGCGGTTCCTGTGCGGGGGCA 1033

Qy     476 TACTCATCAGCTCCTGCTTCTCTGCGCCCACTGCTTCTCCAGGAGAGGTTTCCGC 535
Db    1034 TACTCATCAGCTCCTGCTTCTCTGCGCCCACTGCTTCTCCAGGAGAGGTTTCCGC 1093

Qy     536 CCCACCACCTGACGGTGATCTTGGGCAGAACACTCGGAGCGGCGGCCCCCCAGGCAAACT 595
Db    1094 CCCACCACCTGACGGTGATCTTGGGCAGAACACTCGGAGCGGCGGCCCCCCAGGCAAACT 1153

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACA 655
Db    1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACTTACGACA 1213

Qy     656 ATGACATTGCGCTGCGCAGGTAGTTCGTGCGCCGGAACACTCGGAGACACTCGGGGTG 715
Db    1214 ATGACATTGCGCTGCGCAGGTAGTTCGTGCGCCGGAACACTCGGAGACACTCGGGGTG 1273

Qy     716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGAGACTGCAGCTGACCGGAGTGTG 775
Db    1274 TGGTCCGCACTGTGTGCCTTCCCCCGGCGAGACTGCAGCTGACCGGAGTGTG 1333

Qy     776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
Db    1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1393

Qy     836 AGCTCATGTCAGACTGTACCCATCATCAGCCGCTGCAACACTACCAACTACTACCAACTGAAT 895
Db    1394 AGCTCATGTCAGACTGTACCCATCATCAGCCGCTGCAACACTACCAACTACTACCAACTGAAT 1453

Qy     896 CAGTCACCGACAACATGTCGGCATCATCAGCTGCTGGAGACACTCGGAGCGGCGGCCCCCCAGGCAAACT 955
Db    1454 CAGTCACCGACAACATGTCGGCATCATCAGCTGCTGGAGACACTCGGAGCGGCGGCCCCCCAGGCAAACT 1513

Qy     956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGTCTGAACGATGGCCGCA 1015
Db    1514 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGTCTGAACGATGGCCCGCA 1573

Qy    1016 TGACTTTGGTGGGCATCATCAGCTGTGGGGCCTGGGCTGGGCGTGGGGTG 1075
Db    1574 TGACTTTGGTGGGCATCATCAGCTGTGGGGCCTGGGCTGGGCGTGGGGTG 1633

Qy    1076 TGTACACAAAGGTTACCAACTACTACCTAGACTGAATTCGTGACAACATGGACCG 1128
Db    1634 TGTACACAAAGGTTACCAACTACTACCTAGACTGAATTCGTGACAACATGCGACCG 1686
```

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGAGAGACACTCGGAGCGGCGGGCCCCAGGCAAACT 955
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1454 CAGTCACCGACAACATGCTGTGTGCTGAGAGACACTCGGAGCGGCGGGCCCCAGGCAAACT 1513

Qy   956 TGCACGACGGCCTGCCAGGCGATTCGGGAGGCGGCCCCCTGGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514 TGCACGACGGCCTGCCAGGCGATTCGGGAGGCGGCCCCCTGGTGTCTGAACGATGGCCGCA 1573

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGTGTGGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1574 TGACTTTGGTGGGCATCATCAGCTGGTGTGGACAGAAGGATGTCCCGGGTG 1633

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1634 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCCGACCG 1686

RESULT 11
I08789
LOCUS   I08789   1689 bp   DNA   linear   PAT 02-DEC-1994
DEFINITION   Sequence 3 from Patent WO 8804690.
ACCESSION   I08789
VERSION   I08789.1   GI:588494
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1689)
  AUTHORS   Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.
  JOURNAL   Patent: WO 8804690-A 3 30-JUN-1988;
FEATURES   Location/Qualifiers
     source   1..1689
                     /organism="unknown"
BASE COUNT   379 a   480 c   505 g   325 t
ORIGIN

Query Match   94.4%; Score 1065; DB 6; Length 1689;
Best Local Similarity   99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 673

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCGTCCTCCGTGCAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 GCACGCACAGCCTCACCGAGTCGGGTGCGTCCTCCGTGCAATTCCATGATCCTGA 733

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   734 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 793

Qy   236 ATTACTGCCGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCA 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   794 ATTACTGCCGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCA 853

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTTCGCCTCCACCTGCGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   854 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTTCGCCTCCACCTGCGGCCTGAGACAGTACA 913

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGGCCGACATCGCCTCCCACCCCTGGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   914 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGGCCGCACATCGCCTCCCACCCCTGGC 973

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA 1033

Qy   476 TACTCATCAGCTCCTGGATTCTCTGCCGCCACTGCTTCCAGGAGAGGTTTCCGC 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1034 TACTCATCAGCTCCTGGATTCTCTGCCGCCACTGCTTCCAGGAGAGGTTTCCGC 1093

Qy   536 CCCACCACCTGACGGTGATCTTTGGGCAGAACATACCGGGTGGTGTCCCTGGCGAGGAGAGC 595

Db   1094 CCCACCACCTGACGGTGATCTTTGGGCAGAACATACCGGGTGGTGTCCCTGGCGAGGAGAGC 1153

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1213

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1214 ATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1273

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGCCAGCTGCCAGCTGCCGACTGCAGCGGAGTGTG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274 TGGTCCGCACTGTGTGCCTTCCCCGGCCAGCTGCCAGCTGCCGACTGCAGCGGAGTGTG 1333

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCGGCTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCGGCTGAAGG 1393

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACTTAACAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1394 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACTTAACAGAA 1453

Qy   896 CAGTGACAACATGCTGTGTGCTGAGAGACACTCGGAGCGCGGGCGGGCCCCAGGCAAACT 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454 CAGTGACAACATGCTGTGTGCTGAGAGACACTCGGAGCGCGGGCGGGCCCCAGGCAAACT 1513

Qy   956 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGCGGGCCCCCTGGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGCGGGCCCCCTGGTGTCTGAACGATGGCCGCA 1573

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1574 TGACTTTGGTGGGCATCATCAGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1633

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1634 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1686

RESULT 12
A06611
LOCUS   A06611   1780 bp   DNA   linear   PAT 23-JUN-1993
DEFINITION   Synthetic nucleotide sequence of the 1.78kb t-PA BamHI fragment of
             pt-PABam.
ACCESSION   A06611
VERSION   A06611.1   GI:412906
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1 (bases 1 to 1780)
  AUTHORS   .
  JOURNAL   Patent: WO 9000600-A 11 25-JAN-1990;
FEATURES   Location/Qualifiers
     source   1..1780
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS   17..1705
                     /codon_start=1
                     /transl_table=11
                     /product="t-PA analogue"
                     /protein_id="CAA00577.1"
                     /db_xref="GI:412907"
                     /translation="MDAMKRGLCCVLLLCGAVPVSPSQEIHARFRRGARSYQVICRDE
                     KTQMIYQQHQSWLRPVLRSNRVEYCWCNSGRAQCHSVPVKSCSEPRCFNGGTCQQALY
                     FSDFVCQCPEGFAGKCCEIDTRATCYEDQGISYRGTWSTAESGAECTNWNSSALAQKP
                     YSGRRPDAIRLGLGHNYCRNPDRDSKPWCVVFKAGKYSSEFCSTPACSEGNSDCYFG
                     NGSAYRGTHSLTESGASCLPWNSMILIGKVTAQNPSAQALGLKHNYCRNPDGDAKP
                     WCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRR
                     SPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFEVEKY
                     IVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYG

```
Db   1470  CAGTCACCGACAACATGCTGTGTCTGAGACACTCGGAGCGGGGGCCCCAGGCAAACT  1529
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    956  TGCACGACGCCTGCCAGGGCGATTCGGAGGAGCCGATTCGGAGGAGCCCCGTGTCTGACGATGGCCGCA  1015
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530  TGCACGACGCCTGCCAGGGCGATTCGGAGGAGCCGATTCGGAGGAGCCCCTGTGTCTGACGATGGCCGCA  1589

Qy   1016  TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCGCTGGGCTGTGGACAGAAGGATGTCCCGGGTG  1075
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590  TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCGCTGGGCTGTGGACAGAAGGATGTCCCGGGTG  1649

Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1702
```

RESULT 13
```
A10226
LOCUS       A10226                    1780 bp   DNA   linear   PAT 17-FEB-1997
DEFINITION  Synthetic DNA for mutant tissue plasminogen activator (t-PA).
ACCESSION   A10226
VERSION     A10226.1   GI:490657
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1780)
  AUTHORS   Petersen,L.C. and Boel,E.
  TITLE     A tissue plasminogen activator analogue
  JOURNAL   Patent: EP 0351246-A 7 17-JAN-1990;
            NOVO-NORDISK A/S; NOVO NORDISK A/S
FEATURES             Location/Qualifiers
     source          1..1780
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS             17..1705
                     /codon_start=1
                     /transl_table=11
                     /product="mutant tissue plasminogen activator (t-PA)"
                     /db_xref="CAA00883.1"
                     /db_xref="GI:490658"
                     /translation="MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARSYQVICRDE
                     KTQMIYQQHQSWLRPVLRSNRVEYCWCNSGRAQCHSVPVKSCSEPRCFNGGTCQQALY
                     FSDFVCQCPEGFAGKCCEIDTRATCYEDQGISYRGTWSTAESGAECTNWNSSALAQKP
                     YSGRRPDAIRLGLGNHNYCRNPDRDSKPWCVFVKAGKYSSEFCSTPACSEGNSDCYFG
                     NGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDAKP
                     WCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRR
                     SPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKPEVEKY
                     IVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYG
                     KHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA
                     CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP"
     misc_difference 371..373
                     /note="in t-PA C87S,K419S"
                     /replace="agt"
     misc_difference 1367..1372
                     /note="in t-PA K419S, H420T"
                     /replace="agtact"
     misc_difference 1367..1369
                     /note="in t-PA C87S, K419S"
                     /replace="tcc"
BASE COUNT      405 a    513 c    523 g    339 t
ORIGIN
```

Query Match 94.4%; Score 1065; DB 6; Length 1780;
Best Local Similarity 99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy    56  CCCAGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTTACCGGTG  115
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   630  CCCTGCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGGTG  689

Qy   116  GCACGCACAGCCTCACCGAGTCCGGTGCCTCCTGCCTCCCCGTGGAATTCCATGATCCTGA  175
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   690  GCACGCACAGCCTCACCGAGTCCGGTGCCTCCTGCCTCCCCGTGGAATTCCATGATCCTGA  749
```

```
KEYWORDS     JP 198710457-A/1.
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1  (bases 1 to 1836)
  AUTHORS    Pooru,P.H., Narendaa,K.K. and Shiyouguangu,R.R.
  TITLE      POLYCLINGLE PLASMINOGEN ACTIVATING FACTOR
  JOURNAL    Patent: JP 198710457-A 1 15-MAY-1987;
             AMERICAN HOME PROD CORP
             OS   Artificial gene
             OC   Artificial sequence; Genes.
             OS   homo sapiens
             PN   JP 198710457-A/1
             PD   15-MAY-1987
            ·PF   13-AUG-1986 JP 1986191173
             PR   14-AUG-1985 US 85    766163, 11-JUL-1986 US 86    884835 PI
             POORU POOUEN HANGU, NARENDAA KUMAARU KARUYAN, PI   SHIYOUGUANGU RIN
             RII
             PC   C12N9/52,A61K35/12,A61K35/74,A61K37/02,A61K37/54,C07H21/04, PC
             C12N15/00,
             PC   (C12N9/52,C12R1:19),(C12N15/00,C12R1:19);
             CC   strandedness: Double;
             CC   topology: Linear;
             CC   hypothetical: No;
             CC   anti-sense: No;
             CC   *source: cell_type=fibroblast;
             CC   *source: cell_line=WI-38cell;
             FH   Key                 Location/Qualifiers
             FH
             FT   CDS                 12..1793
             FT                       /product='this-kringle plasminogen activator'
             FT   sig_peptide         12..71
             FT                       /product='signal peptide of urokinase' FT
             FT                 72..1790
             FT                       /product='this-kringle plasminogen activator'.
FEATURES             Location/Qualifiers
     source          1..1836
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   422 a   537 c   528 g   349 t
ORIGIN

Query Match     94.4%;  Score 1065;  DB 6;  Length 1836;
Best Local Similarity  99.5%;  Pred. No. 6.5e-244;
Matches 1068;   Conservative  0;   Mismatches  5;   Indels  0;   Gaps  0;
```

```
Qy      56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 777

Qy     116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCGTGCCTGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCGTGGAATTCCATGATCCTGA 837

Qy     176 TAGGCAAGGTTTACACAGCACAGACACCCAGTGCCAGGCACTGGGCCTGGGCAAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 TAGGCAAGGTTTACACAGCACAGACACCCAGTGCCAGGCACTGGGCCTGGGCAAAACATA 897

Qy     236 ATTACTGCCGGAATCCTGAGGGGATGGGCACAGACTCGGCGTGCCACCTGGTGCCACGTG 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 ATTACTGCCGGAATCCTGATGGGGATGGGCACAGACTCGGCGTGCCACCTGGTGCCACGTG 957

Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTTCGGGTGGGCCTGGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTTCGGGTGGGCCTGGGCCTGAGACAGTACA 1017

Qy     356 GCCAGGCCTCAGTTTCGCATCAAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTCGGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 GCCAGGCCTCAGTTTCGCATCAAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTCGGC 1077

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGGAGGACGGCGTTCGCCGGAGAGGAGGTTTCCGGC 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     690 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGCTCCGTGGAATTCCATGATCCTGA 749

Qy     176 TAGGCAAGGTTTACACAGCACAGACACCCAGTGCCAGGCACTGGGCCTGGGCAAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     750 TAGGCAAGGTTTACACAGCACAGACACCCAGTGCCAGGCACTGGGCCTGGGCAAAACATA 809

Qy     236 ATTACTGCCGGAATCCTGATGGGGATGGGCACAGACTCGGCGTGCCACCTGGTGCCACGTG 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     810 ATTACTGCCGGAATCCTGATGGGGATGGGCACAGACTCGGCGTGCCACCTGGTGCCACGTG 869

Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTTCGGGTGGGCCTGGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     870 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTTCGGGTGGGCCTGGGCCTGAGACAGTACA 929

Qy     356 GCCAGGCCTCAGTTTCGCATCAAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTCGGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     930 GCCAGGCCTCAGTTTCGCATCAAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTCGGC 989

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGGAGGAGCGGTTCGCCCGGAGAGGAGGTTTCCGGC 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     990 AGGCTGCCATCTTTGCCAAGCACAGGAGGAGCGGTTCGCCCGGAGAGGAGGTTTCCGGC 1049

Qy     476 TACTCATCAGCTCCTGCTGGAATCTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCCGC 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1050 TACTCATCAGCTCCTGCTGGAATCTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCCGC 1109

Qy     536 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1110 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGC 1169

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1170 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1229

Qy     656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCGTGCCCAGGAGCAGCG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1230 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCGTGCCCAGGAGCAGCG 1289

Qy     716 TGGTCCGACTGTGTGCCTTCCCCGGGACCTGCAGCTGGACCTGGACGGAGTGTG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1290 TGGTCCGACTGTGTGCCTTCCCCGGGACCTGCAGCTGGACCTGGACGGAGTGTG 1349

Qy     776 AGCTCTCCGGCTACGGCAAGCATGAGGCATTGAGCCGCTTGTCTCCTTCATTCGGAGGCGGCTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1350 AGCTCTCCGGCTACGGCAAGCATGAGGCATTGAGCCGCTTGTCTCCTTCATTCGGAGGCGGCTGAAGG 1409

Qy     836 AGGCTCATGTCAGACTGTACCATCCAGCCGCGTGCACATCCAGACATTTACTTAACAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1410 AGGCTCATGTCAGACTGTACCATCCAGCCGCGTGCACATCCAGACATTTACTTAACAGAA 1469

Qy     896 CAGTCACCGACAACATGCTGTGTCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1470 CAGTCACCGACAACATGCTGTGTCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1529

Qy     956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1530 TGCACGACGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1589

Qy    1016 TGACTTTGGTGGGGCATCATCAGCTGGGGCTGGGCTGTGTGGAACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1590 TGACTTTGGTGGGGCATCATCAGCTGGGGCTGGGCTGTGTGGAACAGAAGGATGTCCCGGGTG 1649

Qy    1076 TGTACACAAAGGTTACCACTACCTAGACTGATTCGTGACACACATCGGACACG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1650 TGTACACAAAGGTTACCACTACCTAGACTGATTCGTGACACACATCGGACACG 1702
```

```
RESULT 14
E01176
LOCUS       E01176          1836 bp    DNA    linear    PAT 29-SEP-1997
DEFINITION  DNA encoding tris-kringle plasminogen activator.
ACCESSION   E01176
VERSION     E01176.1   GI:2169435
```

```
Db  1078 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCA 1137
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   476 TACTCATCAGCTCCTGGATTCTCTGCTCCTCCCACTGCTTCCAGGAGAGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1138 TACTCATCAGCTCCTGCTGGATTCTCTGCTCCGCCCACTGCTTCCAGGAGAGTTTCCGC 1197

Qy   536 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1198 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 1257

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1258 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1317

Qy   656 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTTCGTCCCGCTTGGGGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1318 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTTCGTCCCGCTTGGGGCG 1377

Qy   716 TGGTCCGACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGGACGTGCCGGAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1378 TGGTCCGACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGGACGTGCCGGAGTGTG 1437

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1438 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1497

Qy   836 AGGCTCATGTGGCATCATCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1498 AGGCTCATGTGGCATCATCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1557

Qy   896 CAGTCACCGACAACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCGGCCCCAAACT 955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1558 CAGTCACCGACAACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCGGCCCCAAACT 1617

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTGTCTGAACGATGGCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTGTCTGAACGATGGCCCCA 1677

Qy  1016 TGACTTTGGTGGGCATCATCAGCTACCAACTACCACCTAGACTGGATTCGTGACAACATGCGGACCG 1075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678 TGACTTTGGTGGGCATCATCAGCTACCAACTACCACCTAGACTGGATTCGTGACAACATGCGGACCG 1737

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1738 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG 1790
```

```
Db   718 CCCTGCCTGCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 777
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   116 GCACGCACAGCCTCACCGAGTCGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 GCACGCACCAGCCTCACCGAGTCGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 837

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCTGGGCCTGGGCCTGGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCTGGGCCTGGGCCTGGGCAAACATA 897

Qy   236 ATTACTGCCGGAATCCTGATGGATGCAAGCCTCGTGCCACGTGCTGAAGAACCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 ATTACTGCCGGAATCCTGATGGGATGCAAGCCTCGTGCCACGTGCTGAAGAACCGCA 957

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGGCCTCCACCTGCGGGCCTACGACAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGGCCTCCACCTGCGGGCCTACGACAGTACA 1017

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1018 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 1077

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1078 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCA 1137

Qy   476 TACTCATCAGCTCCTGGATTCTCTGCTCCTCCCACTGCTTCCAGGAGAGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1138 TACTCATCAGCTCCTGCTGGATTCTCTGCTCCGCCCACTGCTTCCAGGAGAGTTTCCGC 1197

Qy   536 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1198 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 1257

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1258 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1317

Qy   656 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTTCGTCCCGCTGCAGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1318 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTTCGTCCCGCTGCAGCG 1377

Qy   716 TGGTCCGACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGGACGTGTGACAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1378 TGGTCCGACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGGACGTGTGACAGTGTG 1437

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1438 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1497

Qy   836 AGGCTCATGTGGCATCATCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1498 AGGCTCATGTGGCATCATCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1557

Qy   896 CAGTCACCGACAACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCGGCCCCAAACT 955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1558 CAGTCACCGACAACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCGGCCCCAAACT 1617

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTGTCTGAACGATGGCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTGTCTGAACGATGGCCCCA 1677

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGATTCGTGACAACATGCGGACCG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678 TGACTTTGGTGGGCATCATCAGCTGGATTCGTGACAACATGCGGACCG 1737

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1738 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG 1790
```

```
RESULT 15
I01581
LOCUS       I01581                  1836 bp ss-DNA     linear   PAT 21-MAY-1993
DEFINITION  Sequence 1 from Patent US 4916071.
ACCESSION   I01581
VERSION     I01581.1  GI:271057
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1836)
AUTHORS     Hung,P.P., Kalyan,N.K. and Lee,S..g.L.
TITLE       Poly-kringle plasminogen activator
JOURNAL     Patent: US 4916071-A 1 10-APR-1990;
            American Home Products Corporation; New York, NY
FEATURES    Location/Qualifiers
     source 1..1836
            /organism="unknown"
BASE COUNT    422 a   537 c   528 g   349 t
ORIGIN

Query Match        94.4%;  Score 1065; DB 6; Length 1836;
Best Local Similarity  99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    56 CCCAGGCGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
```

Search completed: August 19, 2003, 04:46:34

us-09-987-455-5.rge

Job time : 4737.31 secs

THIS PAGE BLANK (USPTO)

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       August 18, 2003, 19:51:53 ; Search time 397.386 Seconds
                                            (without alignments)
                                            7662.485 Million cell updates/sec

Title:            US-09-987-455-5
Perfect score:    1128
Sequence:         1 atgaaaaagacagctatgc........ttcgtgacaacatgcgaccg 1128

Scoring table:    IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched:         252756 seqs, 134719017 residues

Total number of hits satisfying chosen parameters:     5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :  N_Geneseq_19Jun03:*
     1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
     2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
     3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
     4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
     5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
     6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
     7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
     8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
     9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
    10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
    11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
    12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
    13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
    14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
    15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
    16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
    17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
    18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
    19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
    20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
    21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
    22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
    23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
    24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
    25:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1128	100.0	1128	24	AAD40614	OmpA-K2S fusion pr
2	1065	94.4	1314	10	AAN91123	Sequence of coding
3	1065	94.4	1341	11	AAQ06762	Sequence of coding
4	1065	94.4	1419	10	AAN91133	Sequence of coding
5	1065	94.4	1689	9	AAN82178	Tissue plasminogen
6	1065	94.4	1780	11	AAQ01358	Sequence encoding
7	1065	94.4	2100	10	AAN91119	Sequence of native
8	1065	94.4	2100	11	AAQ05532	Plasmid pST112 enc
9	1065	94.4	2162	16	AAO87370	Human tPA clone.
10	1065	94.4	2162	19	AAV37294	Human tissue plasm
11	1065	94.4	2547	7	AAN60659	Sequence encoding
12	1065	94.4	2560	10	AAN90542	Plasmid pKG12 cont
13	1065	94.4	7533	8	AAN70491	Entire sequence of
14	1065	94.4	7533	9	AAN81970	Plasmid pEMp1-tPA.
15	1064	94.3	1680	17	AAT27588	Novel plasminogen
16	1064	94.3	1680	17	AAT27587	Glycosylated plasm
17	1063.4	94.3	1614	11	AAQ03581	Sequence encoding
18	1063.4	94.3	1614	12	AAQ11550	Sequence encoding
19	1063.4	94.3	1659	12	AAQ11551	Tissue type plasmi
20	1063.4	94.3	1686	21	ABL52893	Tissue plasminogen
21	1063.4	94.3	1689	9	AAN82179	Mutated t-PA analo
22	1063.4	94.3	1689	11	AAQ04903	Part of tPA024 gen
23	1063.4	94.3	1689	11	AAQ04904	Part of tPA023 gen
24	1063.4	94.3	1689	11	AAQ05177	Sequence encoding
25	1063.4	94.3	1689	12	AAQ13379	T-PA67+ mutant wit
26	1063.4	94.3	1689	24	ABK52546	Human tissue plasm
27	1063.4	94.3	1740	17	AAT27585	Full-length tissue
28	1063.4	94.3	1955	21	AAZ31998	Human tissue plasm
29	1063.4	94.3	1989	7	AAN60700	Sequence encoding
30	1063.4	94.3	2091	12	AAQ12073	T-PA variant contg
31	1063.4	94.3	2099	11	AAQ05534	Plasmid plgNTQTPA
32	1063.4	94.3	2100	12	AAQ12072	T-PA with -ve char
33	1063.4	94.3	2170	8	AAN70248	Sequence encoding
34	1063.4	94.3	2170	16	AAQ86576	Human tissue plasm
35	1063.4	94.3	2249	11	AAQ05535	Plasmid plgN delta
36	1063.4	94.3	2313	11	AAQ05944	Sequence encoding
37	1063.4	94.3	2457	7	AAN60140	cDNA sequence enco
38	1063.4	94.3	2509	10	AAN90508	cDNA of human tiss
39	1063.4	94.3	2519	24	ABK52545	Human tissue plasm
40	1063.4	94.3	2519	24	ABL59544	Human tissue type
41	1063.4	94.3	2547	4	AAN30001	Sequence encoding
42	1063.4	94.3	7360	17	AAT15930	DHFR/intron (WTras
43	1062	94.1	1065	24	AAD40616	Human kringle 2 se
44	1062	94.1	1068	10	AAN91120	Sequence of coding

```
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

ALIGNMENTS

```
RESULT 1
AAD40614
ID   AAD40614 standard; DNA; 1128 BP.
XX
AC   AAD40614;
XX
DT   30-OCT-2002  (first entry)
XX
DE   OmpA-K2S fusion protein encoding DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cereboprotective; cardiant; ompA; fusion protein; gene; ds.
XX
OS   Unidentified.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1128
FT                  /*tag= a
FT                  /product= "OmpA-K2S fusion protein"
FT                  /transl_except= (pos:1126..1128, aa:Pro-Gly)
FT                  /note= "No stop codon"
FT                  /partial
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
```

```
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25034.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 7; Page 30; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA encoding a fusion protein comprising OmpA and K2S protein.
XX
SQ   Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;

Query Match             100.0%;  Score 1128;  DB 24;  Length 1128;
Best Local Similarity   100.0%;  Pred. No. 5.1e-239;
Matches  1128;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG   60

Qy     61 GCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACG  120

Qy    121 CACAGCCTCACCGAGTCGGGTGCCCTCCTGCGCCTCCTCCGTGGAATTCCATGATCCTGATAGGC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACAGCCTCACCGAGTCGGGTGCCCTCCTGCGCCTCCTCCGTGGAATTCCATGATCCTGATAGGC  180

Qy    181 AAGGTTTACACAGCACAGACAGAACCCCAGTGCCCAGGCACTGGGCACTGGGCAAACATAATTAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAGGTTTACACAGCACAGACAGAACCCCAGTGCCCAGGCACTGGGCACTGGGCAAACATAATTAC  240

Qy    241 TGCCGGAATCCTGATGGGGATGCCAAGCCTGTGCCACGTGCTGAAGAACCGCAGGCTG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGCCGGAATCCTGATGGGGATGCCAAGCCTGTGCCACGTGCTGAAGAACCGCAGGCTG  300

Qy    301 ACGTGGGAGTACTGTGATGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGCAGCCAG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACGTGGGAGTACTGTGATGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGCAGCCAG  360

Qy    361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCT  420

Qy    421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTC  480

Qy    481 ATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCAC  540

Qy    541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGAGGAGGAGCAGAAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGAGGAGGAGCAGAAA  600

Qy    601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC  660

Qy    661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCGCTGTCCCGTGCCCAGGAGCAGCGTGGTC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCGCTGTCCCGTGCCCAGGAGCAGCGTGGTC  720

Qy    721 CGCACTGTGTGCCTTCCCCGGCGGGACCTGCAGCTGCCGGACTGCGGACTGGAGTGTGAGCTC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CGCACTGTGTGCCTTCCCCGGCGGGACCTGCAGCTGCCGGACTGCGGACTGGAGTGTGAGCTC  780

Qy    781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAGGAGGCT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAGGAGGCT  840

Qy    841 CATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTC  900

Qy    901 ACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGCCAAACTTGCAC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGCCAAACTTGCAC  960

Qy    961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACT 1020

Qy   1021 TTGGTGGGCATCATCAGCTGGGGCTGCTGGAGCAGAAGGATGTCCCGGGTGTGTAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTGGTGGGCATCATCAGCTGGGGCTGCTGGAGCAGAAGGATGTCCCGGGTGTGTAC 1080

Qy   1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128


RESULT 2
AAN91123
ID   AAN91123 standard; DNA; 1314 BP.
XX
AC   AAN91123;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of coding region in plasmid pTQKpA delta trp.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key                 Location/Qualifiers
FT   CDS                 1..1065
FT                       /*tag=  a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUISAWA PHARM CO LTD.
XX
```

us-09-987-455-5.rng

', Wed Aug 20 09:29:41 2003

```
PI    Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR    WPI; 1989-040625/06.
DR    P-PSDB; AAP94409.
XX
PT    New tissue plasminogen activator -
PT    comprising finger and growth factor domains lacking tPA for
PT    longer half-life and stronger thrombolytic activity.
XX
PS    Disclosure; Page 7; 68pp; English.
XX
CC    (Updated on 03-OCT-2002 to add missing OS field.)
CC    (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ    Sequence 1314 BP; 286 A; 386 C; 393 G; 249 T; 0 other;

Query Match           94.4%;   Score 1065;  DB 10;  Length 1314;
Best Local Similarity 99.5%;   Pred. No. 3.9e-225;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  5;  Gaps  0;
```

```
Qy     56 CCCAGGCGCCCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
          || ||     |||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 CCCCTGCCTGCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 298

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCGCCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCGCCCGTGGAATTCCATGATCCTGA 358

Qy    176 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCAGGCACTGGGCTGGGCAAACATA 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCAGGCACTGGGCTGGGCAAACATA 418

Qy    236 ATTACTGCCGGAATCCTGATGGGATGGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 ATTACTGCCGGAATCCTGATGGGATGGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 478

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTGCCACCTGCGGCCTGAGACAGTACA 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTGCCACCTGCGGCCTGAGACAGTACA 538

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 598

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    599 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCA 658

Qy    476 TACTCATCAGCTCCTGCTGGATTCTCTCGCGCGGCCACTGCTTCGCAGGAGAGGTTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    659 TACTCATCAGCTCCTGCTGGATTCTCTCGCGCGGCCACTGCTTCGCAGGAGAGGTTTCCGC 718

Qy    536 CCCACCACCTGACGGTGATGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    719 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 778

Qy    596 AGAAATTTGAAGTCGAAAAATACATTT GTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    779 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 838

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTGCCGTGCCCAGGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTGCCGTGCCCAGGAGCAGCG 898

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACCTGGACGTGCCGGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACCTGGACGTGCCGGAGTGTG 958

Qy    776 AGCTCTCCGGCTACGGCACGCATGAGGGCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGG 835
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 AGCTCTCCGGCTACGGCACGCATGAGGGCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGG 1018

Qy    836 AGGCTCATGTGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 AGGCTCATGTGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1078

Qy    896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGGCCCCCAGGCAAACT 1138

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTCGTGTCTGAACGATGGCCGCA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1139 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTCGTGTCTGAACGATGGCCGCA 1198

Qy   1016 TGACTTGTGGGCATCATCAGCTGGGGCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1199 TGACTTGTGGGCATCATCAGCTGGGGCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTG 1258

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACAACATGCGACCG 1128
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1259 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACAACATGCGACCG 1311
```

```
RESULT 3
AAQ06762
ID   AAQ06762 standard; DNA; 1341 BP.
XX
AC   AAQ06762;
XX
DT   09-JAN-2003  (updated)
DT   01-MAR-1991  (first entry)
XX
DE   Sequence encoding non-glycosylated tPA deriv.
XX
KW   Tissue plasminogen activator; thrombolysis; finger domain;
KW   EGF domain;  ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             4..1341
FT                   /tag= a
FT                   /product=tPA deriv.
XX
PN   EP400545-A.
XX
PD   05-DEC-1990.
XX
PF   28-MAY-1990;  90EP-0110096.
XX
PR   14-JUL-1989;  89DE-3923339.
PR   31-MAY-1989;  89DE-3917781.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Stern A, Kohnert U, Rudolph R, Fischer S, Martin U;
XX
DR   WPI; 1990-363094/49.
DR   P-PSDB; AAR08150.
XX
PT   New non-glycosylated form of tissue plasminogen activator - with
PT   thrombolytic activity and long plasma life
XX
PS   Claim 2; page 16; 21pp; German.
XX
CC   This sequence encodes a non-glycosylated tPA deriv. lacking the
CC   finger- and EGF-domains. The polypeptide has a lower clearance
CC   rate (longer half-life) than natural tPA while retaining thrombo-
CC   lytic activity and stimulation by fibrin.
CC   (Updated on 09-JAN-2003 to add missing OS field.)
XX
SQ   Sequence 1341 BP; 293 A; 397 C; 399 G; 252 T; 0 other;

Query Match           94.4%;   Score 1065;  DB 11;  Length 1341;
Best Local Similarity 99.5%;   Pred. No. 3.9e-225;
Matches 1068;  Conservative  0;  Mismatches  0;  Indels  5;  Gaps  0;
```

RESULT 4
AAN91133
ID   AAN91133 standard; DNA; 1419 BP.
XX
AC   AAN91133;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pmTQk112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Location/Qualifiers
FT   Key
FT   CDS   1..1065
FT   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94416.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;

Query Match   94.4%; Score 1065; DB 10; Length 1419;
Best Local Similarity   99.5%; Pred. No.3.9e-225;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   344 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 403

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTGCGGTGCCTCCTCCGTGCCTCCCGTGAATCCCATGATCCTGA 175
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   404 GCACGCACAGCCTCACCGAGTCGGGTGCCTGCGGTGCCTCCTCCGTGCCTCCCGTGAATCCCATGATCCTGA 463

Qy   176 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCACGTGCCAAGCCCTGGGCACTGGGCCTGGGCCTGGGCCAAACATA 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   464 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCACGTGCCAAGCCCTGGGCACTGGGCCTGGGCCTGGGCCAAACATA 523

Qy   236 ATTACTGCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGCCACGTGCTGAAGAACCGCA 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   524 ATTACTGCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGCCACGTGCTGAAGAACCGCA 583

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCTCCTCCACCTGCGGCCTGAGACAGTACA 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   584 GGCTGACGTGGGAGTACTGTGATGTGCCTCCTCCACCTGCGGCCTGAGACAGTACA 643

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415

Qy   416 AGGCTGCCATCTTTGCCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTCGCGGGGCA 475

Qy   476 TACTCATCAGCTCCGCTGGAATTCTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCCGC 535

Qy   536 CCCACCACCTGACCGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGGGAGC 595

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655

Qy   656 ATGACATTGCCGCTGCGAGCTGAAATGGATTCGTCCCGTGCGTGCCCAGGAGCAGCGCG 715

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGACGCGGCGGACGGAGTGTG 775

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGCTCTCCTTTCTATTCGGAGCGGCTGAAGG 835

Qy   836 AGGCTCATGTGCCATCCAGCCGTGCACATCCAGACATTACTTAACGAGAA 895

Qy   1049 AGGCTCATGTGCCATCCAGCCGTGCACATCCAGACATTACTTAACGAGAA 1108

Qy   896 CAGTCACCGACAACATGCTGTGCTGGAGACACTCGGAGCCGCGGGCGCCCAAACT 955

Qy   1109 CAGTCACCGACAACATGCTGTGCTGGAGACACTCGGAGCCGCGGGCGCCCAAACT 1168

Qy   956 TGCACGACGCCTGCGCAGGCGATTCGGGAGGCCCCCGTGTGTGTCTGAACGATGGCCGCA 1015

Qy   1169 TGCACGACGCCTGCGCAGGCGATTCGGGAGGCCCCCGTGTGTGTCTGAACGATGGCCGCA 1228

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGCCTGGGCCGTGTGGGACAGAAGGGATGTCCCGGGTG 1075

Qy   1229 TGACTTTGGTGGGCATCATCAGCTGGGGCGCCTGGGCCGTGTGGGACAGAAGGATGTCCCGGGTG 1288

Qy   1076 TGTACAAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128

Qy   1289 TGTACAAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1341

```
FT                     /product=modified tPA
XX
PN   JP63230083-A.
XX
PD   26-SEP-1988.
XX
PF   20-MAR-1987;   87JP-0064339.
XX
PR   20-MAR-1987;   87JP-0064339.
XX
PA   (EISA ) EISAI CO LTD.
XX
DR   WPI; 1988-311961/44.
DR   P-PSDB; AAP82581.
XX
PT   Modified tissue plasminogen activator - having glycine-183 and serine-186
PT   residues sustd. with serine and threonine.
XX
PS   Disclosure; Page 7; 16pp; Japanese.
XX
CC   One N-glycosylation site, i.e. NSS (117-119) is substituted with
CC   NSM and the N-glycosylation is removed.
CC   Plasmid encoding the modified tPA is 99-6300 and its transformant is
CC   E.coli RR1-Zem 99-6300 (FERM P-9127).
CC   This modified tPA, used to treat thrombosis, is of high quality and
CC   has a longer half life period in blood.
CC   See also AAN82177-N82179.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

Query Match             94.4%;  Score 1065;  DB 9;  Length 1689;
Best Local Similarity   99.5%;  Pred. No. 4e-225;
Matches 1068;  Mismatches   0;  Conservative   0;  Indels   0;  Gaps   0;
```

```
Qy     56 CCCAGGCGGCCTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614 CCCCTGCCTGCTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 673

Qy    116 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    674 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA 733

Qy    176 TAGGCAAGGTTTACACAGCACAGACACCCCAGTGCCAGGCACTGCCAGCACTGCCTGGGCCAGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    734 TAGGCAAGGTTTACACAGCACAGACACCCCAGTGCCAGGCACTGCCAGCACTGCCTGGGCCAGCAAACATA 793

Qy    236 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGAACCCGA 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    794 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGGAACCCGA 853

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCCTCTTCGCCGACATCGCCTCCTCCACCCCTGGC 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    854 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCCTCTTCGCCGACATCGCCTCCTCCACCCCTGGC 913

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCTCCACCCCTGGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    914 GCCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 973

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGACACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974 AGGCTGCCATCTTTGCCAAGCACAGGACACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCCA 1033

Qy    476 TACTCATCAGCTCCTGCTGGATTCTCTGCCAAGCACAGACATTCTCTGCCAAAACATA 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1034 TACTCATCAGCTCCTGCTGGATTCTCTGCCAAGCACAGACATTCTCTGCCAAAACATA 1093

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGTCCCTGGCGAGGAGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGAGC 1153
```

```
Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCGCCGACATCGCTCCCCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    644 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCGCCGACATCGCTCCCCACCCCTGGC 703

Qy    416 AGGGTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 AGGGTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 763

Qy    476 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCGCTTCCGCCCTTCCAGGAGGAGTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCGCTTCCGCCCTTCCAGGAGGAGTTCCGC 823

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGTGTCCCTGGCGAGGAGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    824 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGTGTCCCTGGCGAGGAGAGC 883

Qy    596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCG 1003

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCCGGGCCAGCTGCAGCTGCCGGACTGGACGGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGGTCCGCACTGTGTGCCTTCCCCCGGGCCAGCTGCAGCTGCCGGACTGGACGGAGTGTG 1063

Qy    776 AGCTTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 AGCTCTCCGGCTACGGCAAGCAAGCATGAGGCCTTGTCTCCTTTATTCGGAGCGGCTGAAGG 1123

Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACACAATCATTTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACACAATCATTTACTTAACAGAA 1183

Qy    896 CAGTCACCGGACAACAACATGCTGTGTGCTGCTGAGACACTCGGAGCGGCGGCGGCCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 CAGTCACCGGACAACAACATGCTGTGTGCTGCTGAGACACTCGGAGCGGCGGCGGCCCCCAGGCAAACT 1243

Qy    956 TGCACGAGCGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGGTGTCTGAACGATGGCCGCA 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 TGCACGAGCGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGGTGTCTGAACGATGGCCGCA 1303

Qy   1016 TGACTTTGTGGGCATCATCAGCTGGGGCGCCTGGGCTGTGTGGACAGAAGGAATGGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 TGACTTTGTGGGCATCATCAGCTGGGGCGCCTGGGCTGTGTGGACAGAAGGAATGGTCCCGGGTG 1363

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTCGTGACAACATGCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTCGTGACAACATGCGACCG 1416
```

```
RESULT 5
AAN82178
ID   AAN82178 standard; DNA; 1689 BP.
XX
AC   AAN82178;
XX
DT   25-MAR-2003  (updated)
DT   10-MAR-2003  (updated)
DT   03-NOV-1990  (first entry)
XX
DE   Tissue plasminogen activator with S-119 substd for M.
XX
KW   Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1689
FT                   /*tag= a
```

```
XX
DR   WPI; 1990-016567/03.
DR   P-PSDB; AAR04699.
XX
PT   New tissue plasminogen activator (t-PA) analogue -
PT   with higher fibrin selectivity than native t-pa, useful for
PT   treating infarction, thrombosis and embolism
XX
PS   Disclosure; Fig 5A-D; 24pp; English.
XX
CC   It is modified in the patent by replacing one or more codons specifying
CC   a positively charged amino acid.  The resulting analogues are inserted
CC   into a replicable expression vector which is used to transform or
CC   transfect a host cell which is grown to express a t-PA analogue.  The
CC   t-PA analogue is useful for the treatment of diseases or disorders
CC   associated with the formation of thrombi in blood vessels e.g.
CC   infarctions, thrombosis and embolism.  The analogue, in the 1-chain form,
CC   exhibits the properties of a proenzyme; however on plasmin-catalysed
CC   cleavage of the one-chain form, the activity of the 2-chain form is fully
CC   retained.  Compared to native t-PA, it has a higher fibrin selectivity,
CC   this results in a fibrinogenolytic agent with a higher fibrin selectivity,
CC   as the fibrinogenolytic activity induced by native 1-chain t-PA.
CC   relative to that induced by native t-PA.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;

Query Match            94.4%;  Score 1065;  DB 11;  Length 1780;
Best Local Similarity  99.5%;  Pred. No. 4e-225;
Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy     56 CCCAGGCGGCCTCGAGGGAAACAGTGACTGCTACTTTGGGATGGGTCAGCCTACCGTG    115
          ||||  |  || ||||||||||||||||||||||||| ||||||| |||||||||||||
Db    630 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTGCTACTTTGGGAATGGGTCAGCCTACCGTG 689

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTGCCTCCGTGCCTCCCGTGCCTCCCGTGGGCAACTCCATGATCCTGA  175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 GCACGCACAGCCTCACCGAGTCGGGTGCCTGCCTCCGTGCCTCCCGTGCCTCCCGTGGGCAACTCCATGATCCTGA  749

Qy    176 TAGGCAAGTTTACACAGCACAGCACAGAACCCCAGTGCCCAGTGCCCAGGCACTGGGCCAAACATA  235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 TAGGCAAGTTTACACAGCACAGCACAGAACCCCAGTGCCCAGTGCCCAGGCACTGGGCCAAACATA  809

Qy    236 ATTACTGCCGGAATCCTGATGGGATGGGATGGGCCAAGCCTCGTGCCACGTGCTGAGAACCCGCA  295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 ATTACTGCCGGAATCCTGATGGGATGGGATGGGCCAAGCCTCGTGCCACGTGCTGAGAACCCGCA  869

Qy    296 GGCTGACGTGGGAGTACTGGAGTGATGATGGGCCTCCTGCCCTCGGGCCTGAGACAGTACA  355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    870 GGCTGACGTGGGAGTACTGGAGTGATGATGGGCCTCCTGCCCTCGGGCCTGAGACAGTACA  929

Qy    356 GCCAGCCTCAGTTTGCATCAAAGGAGGGCTCTTGCGCACATCGCGACATCGCCTCCACCCCTGGC  415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    930 GCCAGCCTCAGTTTGCATCAAAGGAGGGCTCTTGCGCACATCGCGACATCGCCTCCACCCCTGGC  989

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCCGAGAGAGCCGGTTCCGTGTGGGGGGCA  475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    990 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCCGAGAGAGCCGGTTCCGTGTGGGGGGCA  1049

Qy    476 TACTCATCAGCTCCTGGATTCTCTCTGCCGCTTCTCCAGGAGGTTTCCGC  535
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 TACTCATCAGCTCCTGGATTCTCTCTGCCGCTTCTCCAGGAGGTTTCCGC  1109

Qy    536 CCCACCACTGACGTGATCTTGGCAGAACATACCGGGTGGTCCTGGGAGGAGGAGC  595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 CCCACCACTGACGTGATCTTGGCAGAACATACCGGGTGGTCCTGGGAGGAGGAGC  1169

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1229

Qy    656 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGGA                       715
          ||||||||||||||||||||||||||||||||||||||
Db   1170 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGGA
```

```
Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1213

Qy    656 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGGA                       715
          ||||||||||||||||||||||||||||||||||||||
Db   1214 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGGA                       1273

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274 TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  1333

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  1393

Qy    836 AGGCTCATGTCAGACTGTACCATCCAGCGCGCTGCACATCACAACATTACTTAACGAA  895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1394 AGGCTCATGTCAGACTGTACCATCCAGCGCGCTGCACATCACAACATTACTTAACGAA  1453

Qy    896 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGCGCGGGCCCCAGGCAAACT  955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGCGCGGGCCCCAGGCAAACT  1513

Qy    956 TGCAGGACGGCCTGCCAGGGGGATTCGGGAGGGCCCCCGTGGTGTGTCTGAACGATGGCCGCA  1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514 TGCAGGACGGCCTGCCAGGGGGATTCGGGAGGGCCCCCGTGGTGTGTCTGAACGATGGCCGCA  1573

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGGGGCCTGGGCTGGGGCCTGGGGTG  1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1574 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGGGGCCTGGGCTGGGGCCTGGGGTG  1633

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1634 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1686
```

```
RESULT 6
AAQ01358
ID   AAQ01358 standard; DNA; 1780 BP.
XX
AC   AAQ01358;
XX
DT   25-MAR-2003 (updated)
DT   08-SEP-1990 (first entry)
XX
DE   Sequence encoding wild type tissue plasminogen activator (t-PA).
XX
KW   Wild type tissue plasminogen activator (t-PA); infarction treatment;
KW   thrombosis treatment; embolism treatment.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..123
FT                   /*tag= a
FT                   /note="synthetic 5' adaptor"
FT   misc_feature    1738..1780
FT                   /*tag= b
FT                   /note="synthetic 3' adaptor"
XX
PN   EP351246-A.
XX
PD   17-JAN-1990.
XX
PF   14-JUL-1989; 89EP-0307194.
XX
PR   15-JUL-1988; 88DK-0003952.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Petersen LC, Boel E;
```

```
PS   Example 29; Fig 21; 68pp; English.
XX
CC   cDNA sequence of native tPA gene is excised from plasmid pST112, and
CC   digested with BamHI and SalI to form plasmid pST118.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match             94.4%;  Score 1065;  DB 10;  Length 2100;
Best Local Similarity   99.5%;  Pred. No. 4.1e-225;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    56  CCCAGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  115
          ||| ||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   638  CCCTGCCTGCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  697

Qy   116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTCCCGTGAATCCATGATCCTGA  175
          ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db   698  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTCCCGTGAAATCCATGATCCTGA  757

Qy   176  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGCACTGGGCCTGGGCAAACATA  235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   758  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGCACTGGGCCTGGGCAAACATA  817

Qy   236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   818  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  877

Qy   296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   878  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  937

Qy   356  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCGTTCTTCGCCGACATCGCCTCCCACCCTGGC  415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   938  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCGTTCTTCGCCGACATCGCCTCCCACCCTGGC  997

Qy   416  AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGTTCCTGTCGCGGGGCA  475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   998  AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCCGTTCCTGTCGCGGGGCA  1057

Qy   476  TACTCATCAGCTCCTGCTGATTCTCTCTGCCAGCCCACTGCTTCTCCAGGAGAGGTTTCCGC  535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1058  TACTCATCAGCTCCTGCTGATTCTCTCTGCCAGCCCACTGCTTCTCCAGGAGAGGTTTCCGC  1117

Qy   536  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGTCCTGGGTCCTGGGGAGC  595
          ||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db  1118  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGTCCTGGGTGGTCCTGGGGAGC  1177

Qy   596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1178  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1237

Qy   656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGGAGCAGCG  715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1238  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGGAGCAGCG  1297

Qy   716  TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACCGGAGTGTG  775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1298  TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACCGGAGTGTG  1357

Qy   776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1358  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  1417

Qy   836  AGGCTCATGTCAGACTGCTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1418  AGGCTCATGTCAGACTGCTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  1477

Qy   896  CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGACACTCGGAGACACTCGGCAAACT  955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1478  CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGACACTCGGCAAACT  1537
```

```
Db  1230  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGGAGCAGCG  1289
Qy   716  TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACCGGAGTGTG  775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1290  TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACCGGAGTGTG  1349
Qy   776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1350  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  1409
Qy   836  AGGCTCATGTCAGACTGCTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1410  AGGCTCATGTCAGACTGCTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  1469
Qy   896  CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT  955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1470  CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT  1529
Qy   956  TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCGTGGTGTCTGAACGATGGCCGCA  1015
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1530  TGCACGAGCGCCTGCCAGGGCGATTCGGGAGGCCCCGTGGTGTCTGAACGATGGCCGCA  1589
Qy  1016  TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGGACGAAGAGGATGTCCCGGGTG  1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1590  TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGGACGAAGAGGATGTCCCGGGTG  1649
Qy  1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGACCG  1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1650  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGACCG  1702
```

```
RESULT 7
AAN91119
ID   AAN91119 standard; DNA; 2100 BP.
XX
AC   AAN91119;
XX
DT   25-MAR-2003  (updated)
DT   03-OCT-2002  (updated)
DT   18-JUN-1990  (first entry)
XX
DE   Sequence of native tPA in plasmid pST112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            25..1710
FT                  /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;  88EP-0112569.
XX
PR   03-AUG-1987;  87GB-0018298.
PR   26-OCT-1987;  87GB-0025052.
PR   13-NOV-1987;  87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94406.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
```

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   698 GCGCGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 757

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGCCTGGCCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   758 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGCCTGGCCAAACATA 817

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   818 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 877

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGTCCACCTGCGGGCCTGAGACAGTACA 355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   878 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGTCCACCTGCGGGCCTGAGACAGTACA 937

Qy   356 GCCAGCCTCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGCACATGCGTCCACCCCTGGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   938 GCCAGCCTCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGCACATGCGTCCACCCCTGGC 997

Qy   416 AGGCTGCCATCTTGCCAAGCACAGCAGGAGTGGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   998 AGGCTGCCATCTTGCCAAGCACAGCAGGAGTGGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1057

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTGCGGCCACTGCTTCTCCAGGAGAGGTTTCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1058 TACTCATCAGCTCCTGCTGGATTCTCTGCCGGCCACTGCTTCTCCAGGAGAGGTTTCCGC 1117

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGGAGC 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1118 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGGAGC 1177

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1178 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1237

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1238 ATGACATTGCGCTGCTGCAGCTGAAATCGATTCGTCCCGCTGTGCCCAGGAGCAGCG 1297

Qy   716 TGGTCCGGACTGTGTGCCTTCCCGGCGCTGCAGCTGCCGGACTGGACGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1298 TGGTCCGGACTGTGTGCCTTCCCGGCGCTGCAGCTGCCGGACTGGACGGAGTGTG 1357

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCCTTCTCATTCGAGGCGGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1358 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCCTTCTCATTCGAGGCGGCTGAAGG 1417

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCGCGTGCACATCACAACATTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1418 AGGCTCATGTCAGACTGTACCATCCAGCGCGTGCACATCACAACATTACTTAACAGAA 1477

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGCCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1478 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGCCAAACT 1537

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCCGTCGTGTGTCTGAACGATGCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1538 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCCGTCGTGTGTCTGAACGATGCCGCA 1597

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCTGTGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1598 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCTGTGACAGAAGGATGTCCCGGGTG 1657

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1710

RESULT 8
AAQ05532
ID   AAQ05532 standard; DNA; 2100 BP.
XX
AC   AAQ05532;
XX
DT   25-MAR-2003 (updated)
DT   11-DEC-1990 (first entry)
XX
DE   Plasmid pSTi12 encoding novel N-terminal for tissue plasminogen
DE   activator (tPA).
XX
KW   Fibrin; clotting; thrombolytic; vascular disease; stroke;
KW   myocardial infarction; heart attack; pulmonary embolism; ds;
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   25..1710
FT   /*tag= a
FT   mat_peptide   130..1710
FT   /*tag= b
XX
PN   EP379890-A.
XX
PD   01-AUG-1990.
XX
PF   10-JAN-1990; 90EP-0100457.
XX
PR   23-JAN-1989; 89GB-0001422.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Satoh S, Suzuki S, Otsuka K, Kusunoki C;
XX
DR   WPI; 1990-232757/31.
DR   P-PSDB; AAR06237.
XX
PT   New tissue plasminogen activator - having N-terminal peptide of
PT   plasminogen linked to tissue plasminogen activator for increased
PT   stability in vivo.
XX
PS   Disclosure; Page ?; ?pp; English.
XX
CC   tPA with N-terminal peptide of plasminogen linked is more stable in
CC   vivo than the native form. It is useful as a thrombolytic agent in
CC   the treatment of vascular diseases eg myocardial infarction,
CC   pulmonary embolism etc.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match   94.4%; Score 1065; DB 11; Length 2100;
Best Local Similarity 99.5%; Pred. No. 4.1e-225;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56   CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   638 CCCCTGCCTGCTGCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 697

RESULT 9
AAQ87370
ID   AAQ87370 standard; DNA; 2162 BP.
XX

```
AC   AAQ87370;
XX
DT   25-MAR-2003 (updated)
DT   19-SEP-1995 (first entry)
XX
DE   Human tPA clone.
XX
KW   Intron; recombination; combinatorial gene; trans-splicing;
KW   gene therapy; polymerase chain reaction; PCR; primer; amplification;
KW   tissue plasminogen activator; tPA; plasmid TPA-KS+; thrombolytic;
KW   ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     82..334
FT                   /*tag= a
FT                   /note= "signal sequence and finger-like domain"
FT                   335..447
FT   CDS             /*tag= b
FT                   /product= EGF-like domain
FT                   448..714
FT   CDS             /*tag= c
FT                   /product= Kringle-1 domain
FT                   715..972
FT   CDS             /*tag= d
FT                   /product= Kringle-2 domain
FT                   973..2162
FT   CDS             /*tag= e
FT                   /product= catalytic domain
XX
PN   WO9507351-A1.
XX
PD   16-MAR-1995.
XX
PF   12-SEP-1994; 94WO-US10146.
XX
PR   10-SEP-1993; 93US-0119512.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Jarrell KA;
XX
DR   WPI; 1995-123425/16.
XX
PT   New intron-mediated recombinant techniques - used for the
PT   generation and selection of novel genes and gene prods. for use
PT   in therapy
XX
PS   Example 4; Page 59-60; 87pp; English.
XX
CC   A cDNA clone of human tissue plasminogen activator (tPA) was
CC   amplified by PCR using the primers given in AAQ87368-69. The
CC   amplified tPA DNA (AAQ87370) was ligated into vector KS+ to
CC   obtain plasmid TPA-KS+. The construct was used in combinatorial
CC   methods involving RNA splicing-mediated shuffling of tPA domains
CC   in plasmid pINVI (AAQ87347) to generate novel tPAs having
CC   improved thrombolytic properties.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match              94.4%;  Score 1065;  DB 16;  Length 2162;
Best Local Similarity    99.5%;  Pred. No. 4.1e-225;
Matches  1068;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    0;

Qy     56 CCCAGGCGGCCCTCGAGGGAAACAGTGACTGACTCTACTTTGGGAATGGCTCAGCCTACCGTG 115
          ||| || ||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||
Db    695 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 754

Qy    116 GCACGCCACAGCCTCACCGAGTCGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||| |||||| |||||||||||||||||||
Db    755 GCACGCCACAGCCTCACCGAGTCGGTGCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA 814
```

```
Qy    176 TAGGCAAGGTTTACACAGCACAGAGACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATA 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    815 TAGGCAAGGTTTACACAGCACAGAGACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATA 874

Qy    236 ATTACTGCCGGAATCCTGATGGGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    875 ATTACTGCCGGAATCCTGATGGGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 934

Qy    296 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCTGCTCACCTGGGCCTGAGACAGTACA 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    935 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCTGCTCACCTGGGCCTGAGACAGTACA 994

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCACGACATCGCCCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    995 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCACGACATCGCCCACCCCTGGC 1054

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1114

Qy    476 TACTCATCAGCTCCTGGTATTCTCTGCCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115 TACTCATCAGCTCCTGGTATTCTCTGCCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 1174

Qy    536 CCCACCACTGACGGTGATCTTGGGCGCAGAACATACGGGTGGTCCCTGGCGAGGAGGAGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175 CCCACCACTGACGGTGATCTTGGGCGCAGAACATACGGGTGGTCCCTGGCGAGGAGGAGC 1234

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1354

Qy    716 TGGTTCCGCACTGTGTGCCTTCCCCGGCCAGTCGCCGACTGCGACTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355 TGGTTCCGCACTGTGTGCCTTCCCCGGCCAGTCGCCGACTGCGACTGGACGGAGTGTG 1414

Qy    776 AGCTCTCCGGCTACGGCAAGCATCGAGGCCTTGTCCTTCCTTTCTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415 AGCTCTCCGGCTACGGCAAGCATCGAGGCCTTGTCCTTCCTTTCTATTCGGAGCGGCTGAAGG 1474

Qy    836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATTCACAACATTACTTAACAGAA 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1475 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATTCACAACATTACTTAACAGAA 1534

Qy    896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGCCAAACT 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1535 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGCCAAACT 1594

Qy    956 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGGCGCCCCTGGTGTCTGAACGATGCCGCA 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1595 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGGCGCCCCTGGTGTCTGAACGATGCCGCA 1654

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCGTGTGCACAGAAGGATGTCCCGGGTG 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1655 TGACTTTGGTGGGCATCATCAGCTGGGGCCGTGTGCACAGAAGGATGTCCCGGGTG 1714

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1767
```

```
RESULT 10
AAV37294
ID   AAV37294 standard; DNA; 2162 BP.
XX
AC   AAV37294;
XX
DT   10-SEP-1998  (first entry)
```

```
                 |||  |||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 CCCCTGCCTGCCTCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  754

Qy   116 GCACGCAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGGAATTCCATGATCCTGA  175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755 GCACGCAGCCTCACCGAGTCGGGTGCCTCCTGCTCCCGTGGAATTCCATGATCCTGA  814

Qy   176 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGCCAAACATA  235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGCCAAACATA  874

Qy   236 ATTACTGCCGGAATCCTGATGGGATGGCAAGCCTGGTGCCACGTGCTGAAGAACCGCA  295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   875 ATTACTGCCGGAATCCTGATGGGATGGCAAGCCTGGTGCCACGTGCTGAAGAACCGCA  934

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGTCTCCACCTGCGGGCCTGAGACAGTACA  355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   935 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGTCTCCACCTGCGGGCCTGAGACAGTACA  994

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 1054

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCCTGTGTGGGGGCA  475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1055 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCCTGTGTGGGGGCA 1114

Qy   476 TACTCATCAGCTCCTGCTGATTCTCTCGCCCACTGCTTCTTCCAGGAGAGGTTTCCGC  535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1115 TACTCATCAGCTCCTGCTGATTCTCTCGCCCACTGCTTCTTCCAGGAGAGGTTTCCGC 1174

Qy   536 CCCACCACTGACGCGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGC  595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1175 CCCACCACTGACGCGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGC 1234

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294

Qy   656 ATGACATTGCGCTGCTGCAGTCGAAATCGGATTCGTCCCGCTGTGTGCCCAGGAGAGCAGCG  715
         |||||||||||||||||||||||||||||||||||||
Db  1295 ATGACATTGCGCTGCTGCAGTCGAAATCGGATTCGTCCTTCATTCGGAGGCGGCTGAAGG 1354

Qy   716 TGGTCCGGACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCGACTGGACGGAGTGTG  775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1355 TGGTCCGGACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCGACTGGACGGAGTGTG 1414

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTGTCCTTCTATTCGGAGCGGCTGAAGG  835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1415 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTGTCCTTCTATTCGGAGCGGCTGAAGG 1474

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGTCACATCACAACATTACTTAACAGAA  895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1475 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGTCACATCACAACATTACTTAACAGAA 1534

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACGTGGCGGGCGGCCCCAGGCAAACT  955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1535 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGGCGGCGGCCCCAGGCAAACT 1594

Qy   956 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGAGGCCCCCTGTGTTGTCTGACGATGCCGCCA 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1595 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGAGGCCCCCTGTGTCTGACGATGCCGCCA 1654

Qy  1016 TGACTTTCGTGGGGCATCATCAGCTGGGCTGGGCTGGGCGTGTGACAGAAGGATGTCCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1655 TGACTTTCGTGGGGCATCATCAGCTGGGCTGGGCGTGTGACAGAAGGATGTCCCGGGTG 1714

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1767
```

XX
DE   Human tissue plasminogen activator gene sequence.
XX
KW   Plasmid pINVI; reverse-splicing intron; group II intron;
KW   exon binding site; domain V motif; branch site acceptor;
KW   nucleophilic group; transesterification; phosphodiester bond;
KW   autocatalytic Y-branched intron; reverse splicing reaction; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   82..334
FT   /*tag= a
FT   misc_feature   335..447
FT   /note= "signal sequence and finger-like domain"
FT   misc_feature   /*tag= b
FT   /note= "EGF-like domain"
FT   misc_feature   448..714
FT   /*tag= c
FT   /note= "Kringle-1 domain "
FT   misc_feature   715..972
FT   /*tag= d
FT   /note= "Kringle-2 domain"
FT   misc_feature   973..2162
FT   /*tag= e
FT   /note= "catalytic domain"
XX
PN   US5780272-A.
XX
PD   14-JUL-1998.
XX
PF   07-JUN-1995; 95US-0488015.
XX
PR   10-SEP-1993; 93US-0119512.
XX
PA   (HARD) HARVARD COLLEGE.
XX
PI   Jarrell KA;
XX
DR   WPI; 1998-413060/35.
XX
PT   Reverse splicing construct containing fragments of autocatalytic
PT   introns - able to cleave and ligate discontinuous nucleic acid for
PT   generating new genes and e.g. ribozymes, libraries of enzymes and
PT   antibodies
XX
PS   Example 4; Columns 53-56; 56pp; English.
XX
CC   The present sequence represents the human tissue plasminogen activator
CC   gene. It was used to construct plasmid TPA-KS+, which is used in the
CC   course of the invention. The specification describes a purified
CC   reverse-splicing intron which comprises a segment comprising a
CC   5'-part of a group II intron, including an exon binding site not
CC   naturally present in the intron and a second segment comprising a
CC   3'-part of a group II intron, including a domain V motif, a branch
CC   site acceptor, and a nucleophilic group for transesterifying a
CC   phosphodiester bond of an RNA. Together the two segments form an
CC   autocatalytic Y-branched intron which catalyses integration of at least
CC   the first segment into substrate RNA by a reverse splicing reaction
CC   The reverse-splicing introns are used, by specific cleavage and ligation
CC   of discontinuous nucleic acid, to generate new genes and gene products,
CC   e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC   manipulation, e.g. replacements for restriction enzymes) or
CC   immunologically active or signal-transducing proteins such as antibody
CC   and enzyme libraries.
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match   94.4%; Score 1065; DB 19; Length 2162;
Best Local Similarity 99.5%; Pred. No. 4.1e-225;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115

RESULT 11

```
AAN60659
ID   AAN60659 standard; DNA; 2547 BP.
XX
AC   AAN60659;
XX
DT   25-MAR-2003 (updated)
DT   22-JUL-1991 (first entry)
XX
DE   Sequence encoding human pre-tissue plasminogen activator (pre-t-PA).
XX
KW   Plasminogen conversion; vascular disease therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             87..191
FT                   /*tag= a
FT   mat_peptide     192..1775
FT                   /*tag= b
XX
PN   GB2173804-A.
XX
PD   22-OCT-1986.
XX
PF   21-APR-1986;  86GB-0609683.
XX
PR   22-APR-1985;  85US-0725468.
PR   01-APR-1986;  86US-0846697.
PR   22-APR-1985;  85US-0725468.
PR   01-APR-1986;  86US-0846697.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Heyneker HL, Vehar GA;
XX
DR   WPI; 1986-280715/43.
DR   P-PSDB; AAP60790.
XX
PT   New mutant forms of human tissue plasminogen activator- having
PT   higher specific activity and resistance to conversion to two chain
PT   form
XX
PS   Example; Fig 2; 34pp; English.
XX
CC   The patentors claim a novel recombinant human t-PA which is
CC   resistant to specific enzymatic cleavage because it is stabilised by
CC   site-directed mutagenesis at a 2-chain cleavage site. Partic. the
CC   natural Arg at position 275 is replaced by Gly or Glu, or Ile at
CC   position 276 is replaced (275 and 276 refer to the posn. of the AAs
CC   in the mature protein; i.e. AAs 310 and 311 of AAP60790). Also new are
CC   (1) DNA sequences; (2) expression vectors; and (3) microorganisms
CC   and cell cultures transformed with these vectors.
CC   (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;

Query Match             94.4%;  Score 1065;  DB 7;  Length 2547;
Best Local Similarity   99.5%;  Pred. No. 4.2e-225;
Matches  1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy   56  CCCAGGCGGCCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGGTCAGCCTACCGTG 115
         |||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  700  CCCCTGCTGCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGGTCAGCCTACCGTG 759

Qy  116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  760  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA 819

Qy  176  TAGGCAAGGTTTACACAGCACAGACACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  820  TAGGCAAGGTTTACACAGCACAGACACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 879

Qy  236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCCTGGTGCCACGTGCTGAAGAACCCCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  880  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCCTGGTGCCACGTGCTGAAGAACCGCA 939

Qy  296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  940  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 999

Qy  356  GCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCGCCGACATGGCCTCCACCCCTGGC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1000  GCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCGCCGACATGGCCTCCACCCCTGGC 1059

Qy  416  AGGCTGCCATCTTTGCCAAGCACAGAGGAGGTCGCCCGGAGGCGGTTCCTGTGCGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1060  AGGCTGCCATCTTTGCCAAGCACAGAGGAGGTCGCCCGGAGGCGGTTCCTGTGCGGGGGCA 1119

Qy  476  TACTCATCAGCTCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCTCCAGGAGAGGTTCCGC 535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1120  TACTCATCAGCTCCTGCTGGATTCTCTCTGCCCGCCCACTGCTTCTCCAGGAGAGGTTCCGC 1179

Qy  536  CCCACCACCTGACGTGATCTTGGGCAGAAACATACCGGGGTGGTCCCTGGCCGGAGGAGC 595
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1180  CCCACCACCTGACGTGATCTTGGGCAGAAACATACCGGGGTGGTCCCTGGCCGGAGGAGC 1239

Qy  596  AGAAATTTGAAGTGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1240  AGAAATTTGAAGTGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1299

Qy  656  ATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1300  ATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCCGTGTGCCCAGGAGAGCAGCG 1359

Qy  716  TGGTTCCGACTGTGTCCTTCCCCCGGCGACCTGCAGCTGCCGACTGGCCGACGGAGTGTG 775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1360  TGGTTCCGCACTGTGTCCTTCCCCCGGCGACCTGCAGCTGCCGACTGGCCGACGGAGTGTG 1419

Qy  776  AGCTCCCGGCCTACGGCAAGCATGAGCCCTGTCCCTTTCTATTCGGAGCGGCTGAAGG 835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1420  AGCTCTCCGGCCTACGGCAAGCATGAGCCCTGTCCCTTTCTGTCCCTTTCTATTCGGAGCGGCTGAAGG 1479

Qy  836  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGGAGACATTCACAACATTACTTAACAGAA 895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1480  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGGAGACATTCACAACATTACTTAACAGAA 1539

Qy  896  CAGTCCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCGGCCCCAGGCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1540  CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCGGCCCCAGGCAAACT 1599

Qy  956  TGCACGACGGCTGCCAGGGCGATTCGGAGGCGCCCCTGGTGTGTCGAACGATGGCCGCA 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1600  TGCACGACGGCTGCCAGGGCGATTCGGAGGCGCCCCTGGTGTGTCGAACGATGGCCGCA 1659

Qy 1016  TGACTTTGGTGGGCATCATCAGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1660  TGACTTTGGTGGGCATCATCAGCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1719

Qy 1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTCGTCGACAACATGCGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1720  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTCGTCGACAACATGCGACCG 1772
```

```
RESULT 12
AAN90542
ID   AAN90542 standard; cDNA; 2560 BP.
XX
AC   AAN90542;
XX
DT   25-MAR-2003 (updated)
DT   05-JUN-1990 (first entry)
XX
DE   Plasmid pKG12 contg. DNA encoding human melanoma t-PA.
XX
KW   Human tissue plasminogen-activator gene; pKG12; ss.
XX
```

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     102..206
FT                   /*tag= a
FT   mat_peptide     ..1787
FT                   /*tag= b
FT                   /product=t-PA
XX
PN   EP297066-A.
XX
PD   28-DEC-1988.
XX
PF   14-JUN-1988;   8BEP-0850207.
XX
PR   18-JUN-1987;   87SE-0002562.
XX
PA   (KABI ) KABIGEN AB.
XX
PI   Pohl G, Hansson L, Loewenadler B;
XX
DR   WPI; 1989-001503/01.
DR   P-PSDB; AAP93716.
XX
PT   Modified tissue plasminogen-activator - having domains deleted and amino
PT   acid changes to increase biological half-life and reduce inactivation.
XX
PS   Disclosure; Fig 1; 18pp; English.
XX
CC   The sequence was obtained from cDNA prepd. from mRNA extracted from
CC   Bowes melanoma cells. The cDNA was used to construct a gene library which
CC   was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated
CC   and shown to comprise the whole coding region for human t-PA as well as
CC   102bp 5'flanking, 760 bp 3' flanking DNA and a poly A tail.
CC   See also AAN91608.
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;

 Query Match             94.4%;  Score 1065;  DB 10;  Length 2560;
 Best Local Similarity   99.5%;  Pred. No. 4.2e-225;
 Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy        56 CCCAGGCGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTG 115
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       715 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTG 774

Qy       116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCCGTGGAATTCCATGATCCTGA 175
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       775 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCCGTGGAATTCCATGATCCTGA 834

Qy       176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGCCTCCAGGCACTGGCCAAACATA 235
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       835 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGCCTCCAGGCACTGGCCAAACATA 894

Qy       236 ATTACTGCCGGAATCCTGATGGGAATGGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       895 ATTACTGCCGGAATCCTGATGGGAATGGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 954

Qy       296 GGGCTGACGTGGGAGTACTGTGATGTGCCCCTCCTGCTCCCTCCACCTGGCGGCTGCAGTACA 355
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       955 GGGCTGACGTGGGAGTACTGTGATGTGCCCCTCCTGCTCCCTCCACCTGGCGGCTGCAGTACA 1014

Qy       356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCACATCGCCTCCCACCCCTGGC 415
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1015 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCACATCGCCTCCCACCCCTGGC 1074

Qy       416 AGGCTGCCATCTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1075 AGGCTGCCATCTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1134

Qy       476 TACTCATCAGCTCCTGCTGGAATTCTCTGCGCCCCACTGCTTCCAGGAGAGTTTCCGC 535
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1135 TACTCATCAGCTCCTGCTGGAATTCTCTGCGCCCCACTGCTTCCAGGAGAGTTTCCGC 1194

Qy       536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGAGC 595
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1195 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGAGC 1254

Qy       596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1255 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1314

Qy       656 ATGACATTGCGCTGCTGCAAATCGGATTCGTCCGCTGTGCCCAGGAGAGCAGCG 715
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1315 ATGACATTGCGCTGCTGCGAATCGGATTCGTCCCGTGTGCCCAGGAGAGCAGCG 1374

Qy       716 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACTGCAGCTGCGACTGGAACGGAGTGTG 775
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1375 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACTGCAGCTGCGACTGGAACGGAGTGTG 1434

Qy       776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1435 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1494

Qy       836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACACATTTACTTAACAGAA 895
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1495 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACACATTTACTTAACAGAA 1554

Qy       896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCAGGCAAACT 955
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1555 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCAGGCAAACT 1614

Qy       956 TGCACGACGGCCTGCCAGGGCGATTCGGAGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1015
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1615 TGCACGACGGCCTGCCAGGGCGATTCGGAGGCCCCCGTGTGTCTGAACGATGGCCGCA 1674

Qy      1016 TGACTTTGGTGGGGCATCATCAGCTGGGGCTGGGGCTGTGGACAGAAGGATGTCCGGGGTG 1075
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1675 TGACTTTGGTGGGGCATCATCAGCTGGGGCTGGGGCTGTGTGGACAGAAGGATGTCCGGGGTG 1734

Qy      1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCG 1128
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1735 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCG 1787
```

```
RESULT 13
AAN70491
ID   AAN70491 standard; cDNA; 7533 BP.
XX
AC   AAN70491;
XX
DT   25-MAR-2003 (updated)
DT   31-OCT-2002 (updated)
DT   07-JUN-1991 (first entry)
XX
DE   Entire sequence of high level expression vector pEMpl-tpa for tissue
DE   plasminogen activator (TPA).
XX
KW   Enhancer-blocking element; expression vector; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   mRNA            20..1708
FT                   /*tag= a
FT   misc_feature    7..1748
FT                   /*tag= b
FT                   /label= TPA cDNA
FT   3'UTR           1749..1955
FT                   /*tag= c
FT                   /note= "from immunoglobulin kappa light chain gene"
FT   3'clip          1956..2164
FT                   /*tag= d
FT   misc_feature    2165..4463
FT                   /note= "as above"
```

```
FT        /tag= e
FT        /label= pBR322
FT        /note= "contg. the ampicillinase gene and the
FT               bacterial origin of replication"
FT  misc_feature  complement (4464..6436)
FT        /tag= f
FT        /label= E.coli gpt gene
FT  promoter  6437..6547
FT        /tag= g
FT        /label= SV40 origin region
FT  misc_signal  6454..6468
FT        /tag= h
FT        /label= SV40 (gpt) CAP site
FT  enhancer  6548..6715
FT        /tag= i
FT        /label= SV40 enhancer (T2 bp repeat)
FT  promoter  6716..6971
FT        /tag= j
FT        /label= lambda-1
FT        /note= "functions as a blocking element"
FT  enhancer  6972..7281
FT        /tag= k
FT        /note= "from immunoglobulin heavy chain (IgH)"
FT  promoter  7282..7466
FT        /tag= l
FT        /note= "from metallothionine 1 (MT)"
FT  misc_signal  7467..7533
FT        /tag= m
FT        /label= MT CAP site
FT  polyA_signal  1930..1935
FT        /tag= n
XX
PN   EP237157-A.
XX
PD   16-SEP-1987.
XX
PF   27-JAN-1987;  87EP-0300658.
XX
PR   07-MAR-1986;  86US-0837595.
XX
PA   (DAMO-) DAMON BIOTECH INC.
PA   (ABBO ) ABBOTT BIOTECH INC.
XX
PI   Gillies SD;
XX
DR   WPI; 1987-258421/37.
XX
PT   DNA construct for high level expression of protein - having a
PT   blocking element so that marker protein is produced only at
PT   levels required for selection
XX
PS   Disclosure; Fig. 9A-D; 20pp; English.
XX
CC   Vector pEMp1-tpa is an example of a vector of the invention. Pref.
CC   the blocking element is a promoter sequence oriented with its native
CC   3' end disposed proximal to the enhancer element and its 5' end
CC   disposed distal to the enhancer element. Alternatively the
CC   interposed blocking element comprises a promoter sequence with its 5'
CC   end disposed proximal to the enhancer element and, at its 3' end, a
CC   gene encoding for a different protein of interest, (see FT). The
CC   vector pEMp1 was constructed by inserting an enhancer-blocking
CC   element into pEM1.
CC   TPA cDNA fragment was inserted into the unique XhoI site present in
CC   both pEM1 and pEMp1. Of 26 transformants obtd. with pEM1-tpa, 5
CC   produced TPA at levels of 140-500 IU/ml. Of 16 transformants obtd.
CC   with pEMp1-tpa, 7 produced TPA at levels of 1000-6000 IU/ml.
CC   (Updated on 31-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 7533 BP; 1921 A; 1899 C; 1844 G; 1868 T; 1 other;
```

Query Match   94.4%; Score 1065; DB 8; Length 7533;
Best Local Similarity   99.5%; Pred. No. 4.8e-225;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        633 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  692

Qy        116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGA  175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGA  752

Qy        176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  812

Qy        236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  872

Qy        296 GGCTGACGTGGGAGTACTGCATGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 GGCTGACGTGGGAGTACTGTGATGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  932

Qy        356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCCGACATGCCTCCCACCCCTGGC  415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCCGACATGCCTCCCACCCCTGGC  992

Qy        416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCGTCTGTGCGGGGCA  475
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        993 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCGTCTGTGCGGGGCA  1052

Qy        476 TACTCATCAGCTCCTGCTGGATTCTTCTCTGCCGCGCCCACTGCTTCCAGGAGAGTTTCCGC  535
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1053 TACTCATCAGCTCCTGCTGGATTCTTCTCTGCCGCGCCCACTGCTTCCAGGAGAGTTTCCGC  1112

Qy        536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGC  595
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1113 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGC  1172

Qy        596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1173 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1232

Qy        656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG  715
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1233 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG  1292

Qy        716 TGGTCCGACACTGTGTGCCTTCCCCGCGGCGACCTGCAGCTGCGACCTGGACCGGAGTGTG  775
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1293 TGGTCCGACACTGTGTGCCTTCCCCGCGGCGACCTGCAGCTGCGACCTGGACCGGAGTGTG  1352

Qy        776 AGCTCTCCGGCTACGGCAAGCATGAGGGCTGTCCTCCTTTCTATTCGGAGGCGGCTGAAGG  835
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1353 AGCTCTCCGGCTACGGCAAGCATGAGGGCTGTCCTCCTTTCTATTCGGAGGCGGCTGAAGG  1412

Qy        836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGAA  895
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1413 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACACATTACTTAACAGAA  1472

Qy        896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACT  955
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1473 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACT  1532

Qy        956 TGCACGACGCCTGCCAGGCGGATTCGGAGGCCCCTGTGTCTGAACGATGGCCGCA  1015
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1533 TGCACGACGCCTGCCAGGCGGATTCGGAGGCCCCTGTGTCTGAACGATGGCCGCA  1592

Qy       1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG  1075
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1593 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG  1652

Qy       1076 TGTTACACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACTGACAACATGCGACCG  1128
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1653 TGTTACACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACTGACAACATGCGACCG  1705
```

us-09-987-455-5.rng

```
RESULT 14
AAN81970
ID   AAN81970 standard; DNA; 7533 BP.
XX
AC   AAN81970;
XX
DT   12-OCT-1990 (first entry)
XX
DE   Plasmid pEMp1-tPA.
XX
KW   Tissue plasminogen activator; myocardial infarction; thrombosis;
KW   embolism; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_RNA        7..1748
FT                   /*tag= a
FT                   /label=tPA_cDNA
FT   CDS             20..1708
FT                   /*tag= b
FT                   /product=t-PA
FT   3'UTR           1749..1955
FT                   /*tag= c
FT                   /label=K_3'UT
FT                   /note="from Ig kappa light chain gene"
FT   polyA_signal    1930..1935
FT                   /*tag= d
FT   misc_RNA        1956..2104
FT                   /*tag= e
FT                   /note="kappa 3 flanking sequence"
FT   misc_RNA        2105..4463
FT                   /*tag= f
FT                   /label=pBR322
FT                   /note="contg. ampicillinase gene and origin of
FT                   replication"
FT   misc_RNA        4464..6971
FT                   /*tag= g
FT                   /label=pSV2-gpt
FT                   /note="non-coding strand"
FT   protein_bind    6437..6547
FT                   /*tag= h
FT                   /label=Cap_site
FT                   /note="from SV40(gpt)"
FT   misc_RNA        6453..6469
FT                   /*tag= i
FT                   /label=origin.region
FT                   /note="from SV40"
FT   repeat_unit     6548..6619
FT                   /*tag= j
FT                   /rpt_type=direct
FT   repeat_unit     6620..6691
FT                   /*tag= k
FT                   /rpt_type=direct
FT   enhancer        6633..6650
FT                   /*tag= l
FT                   /note="from SV40"
FT   misc_RNA        6716..6945
FT                   /*tag= m
FT                   /label=blocking element
FT   protein_bind    6946..6971
FT                   /*tag= n
FT                   /label=cap_site
FT                   /note="from lambda promoter"
FT   enhancer        6972..7281
FT                   /*tag= o
FT                   /note="lambda mRNA"
FT   promoter        7282..7466
FT                   /*tag= p
FT                   /note="from IgH"
FT   protein_bind    7467..7533
FT                   /*tag= q
FT                   /label=cap_site
FT                   /note="from metallothioneine I"
XX
PN   WO8800242-A.
XX
PD   14-JAN-1988.
XX
PF   25-JUN-1987;   87WO-US01569.
XX
PR   26-JUN-1986;   86US-0879038.
XX
PA   (DAMO-) DAMON BIOTECH INC.
XX
PI   Gillies SD;
XX
DR   WPI; 1988-021592/03.
DR   P-PSDB; AAP81913.
XX
PT   Prodn. of human tissue plasminogen activator - by myeloma cells,
PT   pref. in presence of epsilon amino caproic acid.
XX
PS   Disclosure; Page 7; p; English.
XX
CC   The plasmid was prepd by cleaving the very long 3' UT of t-PA cDNA
CC   ( which causes mRNA instability) 34 nucleotides downstream of stop
CC   codon and inserting it intoa pEMp1 vector. The pEMp1 vector was
CC   constructed from the following components (see feature table):
CC   (a) a 2.25 PvuII-BamHI fragment from pSV2-gpt contg. the SV40
CC   enhancer and early region promoter, the E.coli gpt gene, the SV40
CC   small tumour antigen intervening sequence, and the SV40 termination
CC   and polyadenylation signals;
CC   (b) a 2.3 kb PvuII-EcoRI fragment from pBR322 contg. the
CC   ampicillinase gene and the bacterial origin of replication;
CC   (c) a 0.3 kb PvuII-EcoRI fragment contg. an Ig heavy chain enhancer;
CC   (d) a 0.25 kb SacI-BglII fragment contg. the metallothioneine I rromoter;
CC   (e) a 0.4 kb AvaII-HaeIII fragment from the 3' UT of Ig kappa light
CC   chain gene; and
CC   (f) a 0.26 kb xbaI-bstNI fragment contg. the Ig light chain promoter,
CC   TATAA sequence, and the initiation site.
CC   The expression plasmid can be used to transform host cells esp.
CC   J5581, ATCC CRL 9132.
XX
SQ   Sequence 7533 BP; 1918 A; 1892 C; 1847 G; 1875 T; 1 other;
```

```
Query Match             94.4%;  Score 1065;  DB 9;  Length 7533;
Best Local Similarity   99.5%;  Pred. No. 4.8e-225;
Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy         56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 692

Qy        116 GCACGCACAGCCTCACCGAGTGCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGA 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 GCACGCACAGCCTCACCGAGTGCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGA 752

Qy        176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACTGGGCAAACATA 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACTGGGCAAACATA 812

Qy        236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGGTGCCACGTGCTGCTGAAGAACCGCA 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGGTGCCACGTGCTGCTGAAGAACCGCA 872

Qy        296 GGCTGACGTGGGAGTACTGTGATGTGCCCCTCCGTCCACCTGGGCGCCTGGGCCGAGACAGTACA 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 GGCTGACGTGGGAGTACTGTGATGTGCCCCTCCGTCCACCTGGGCGCCTGGGCCGAGACAGTACA 932

Qy        356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGC 415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCCACCCCTGGC 992
```

```
Qy       416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       993 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1052

Qy       476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1053 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 1112

Qy       536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1113 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1172

Qy       596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1173 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1232

Qy       656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1233 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1292

Qy       716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1293 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1352

Qy       776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1353 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1412

Qy       836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1413 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1472

Qy       896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACT 955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1473 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACT 1532

Qy       956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1533 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1592

Qy      1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1593 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1652

Qy      1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1653 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1705

RESULT 15
AAT27588
ID   AAT27588 standard; DNA; 1680 BP.
XX
AC   AAT27588;
XX
DT   25-MAR-2003  (updated)
DT   06-AUG-1996  (first entry)
XX
DE   Novel plasminogen activator DNA.
XX
KW   Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
KW   fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
KW   protein engineering; kringle; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     1..105
FT                   /*tag=  a
FT   mat_peptide     106..1677
FT                   /*tag=  b
XX
PN   US5504001-A.
XX
PD   02-APR-1996.
XX
PF   06-JUN-1994;   94US-0254485.
XX
PR   25-NOV-1987;   87US-0125629.
PR   28-JAN-1992;   92US-0827587.
PR   06-JUN-1994;   94US-0254485.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Foster DC;
XX
DR   WPI; 1996-187699/19.
DR   P-PSDB; AAR96223.
XX
PT   Hybrid plasminogen activator comprises human tPA activator and
PT   N-terminal crosslinking domain from alpha2-plasmin inhibitor  -
PT   useful to treat thrombosis and image blood clots
XX
PS   Example 3; Fig 11; 35pp; English.
XX
CC   A DNA construct (AAT27586) codes for a novel plasminogen
CC   activator (AAR96222) in which the kringle K1 domain (AAR96221)
CC   of plasminogen replaces the native K1 domain of tissue
CC   plasminogen activator (tPA) (AAR96220).  It was obtd. by
CC   inserting a synthetic plasminogen K1 domain DNA (AAT27586)
CC   into tPA cDNA and can be expressed in transformed Escherichia
CC   coli RRI cells (FERM P-9272).  Novel plasminogen activators are
CC   produced that show increased clot lysing specificity or plasma
CC   half-life.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 1680 BP; 387 A; 471 C; 479 G; 343 T; 0 other;

  Query Match             94.3%;  Score 1064;  DB 17;  Length 1680;
  Best Local Similarity   99.5%;  Pred. No. 6.6e-225;
  Matches 1067;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        57 CCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGG 116
             || ||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       606 CCTGGAATGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGG 665

Qy       117 CACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGAT 176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       666 CACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGAT 725

Qy       177 AGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAA 236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       726 AGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAA 785

Qy       237 TTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAG 296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       786 TTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAG 845

Qy       297 GCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAG 356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       846 GCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAG 905

Qy       357 CCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCA 416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       906 CCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCA 965

Qy       417 GGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCAT 476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       966 GGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCAT 1025

Qy       477 ACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCC 536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1026 ACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCC 1085

Qy       537 CCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCA 596
```

```
Db   1086 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCA 1145

Qy    597 GAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAA  656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1146 GAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAA 1205

Qy    657 TGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCAGGAGCAGCGT  716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1206 TGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCAGGAGCAGCGT 1265

Qy    717 GGTCCGACTGTGTGCCTTCCCCGGGCGGACCTGCCAGCTGCCGGACTGGACGGAGTGTGA  776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1266 GGTCCGACTGTGTGCCTTCCCCGGGCGGACCTGCCAGCTGCCGGACTGGACGGAGTGTGA 1325

Qy    777 GCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTTCCTTCTATTCGGAGCGGCTGAAGGA  836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1326 GCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTTCCTTCTATTCGGAGCGGCTGAAGGA 1385

Qy    837 GGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAAC  896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1386 GGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAAC 1445

Qy    897 AGTCACCGACAACATGCTGTGTCTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTT  956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1446 AGTCACCGACAACATGCTGTGTCTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTT 1505

Qy    957 GCACGACGCCTGCCAGGGCGATTCGGAGGAGGCCCCCTGGTGTCTGAACGATGGCCCGCAT 1016
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1506 GCACGACGCCTGCCAGGGCGATTCGGAGGAGGCCCCCTGGTGTCTGAACGATGGCCCGCAT 1565

Qy   1017 GACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGT 1076
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1566 GACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTGT 1625

Qy   1077 GTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1626 GTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCACCG 1677
```

Search completed: August 18, 2003, 23:28:52
Job time : 398.386 secs

```
          GenCore version 5.1.6
     Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     August 18, 2003, 22:51:23 ; Search time 88.2804 Seconds
                                         (without alignments)
                                         5639.757 Million cell updates/sec

Title:            US-09-987-455-5
Perfect score:    1128
Sequence:         1 atgaaaagacagctatcgc........ttcgtgacaacatgcgaccg 1128

Scoring table:    IDENTITY NUC
                  Gapop 10.0 , Gapext 1.0

Searched:         569978 seqs, 22069156 residues

Total number of hits satisfying chosen parameters:    1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        Issued_Patents NA:*
                  1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                  2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                  3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                  4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                  5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                  6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1065	94.4	1419	2	US-08-811-949-62	Sequence 62, Appl
2	1065	94.4	1848	2	US-08-814-412-10	Sequence 10, Appl
3	1065	94.4	2101	2	US-08-811-949-42	Sequence 42, Appl
4	1065	94.4	2162	1	US-08-119-512-3	Sequence 3, Appli
5	1065	94.4	2162	1	US-08-488-015B-3	Sequence 3, Appli
6	1065	94.4	2162	1	US-08-488-015B-25	Sequence 25, Appl
7	1063.4	94.3	1314	2	US-08-811-949-48	Sequence 48, Appl
8	1063.4	94.3	1738	6	520340-1	Patent No. 5200340
9	1063.4	94.3	1955	2	US-08-883-795A-39	Sequence 39, Appl
10	1063.4	94.3	2457	6	5344773-1	Patent No. 5344773
11	1063.4	94.3	7360	1	US-08-286-740-1	Sequence 1, Appli
12	1063.4	94.3	7360	5	PCT-US95-09576-1	Sequence 1, Appli
13	1062	94.1	1068	2	US-08-811-949-44	Sequence 44, Appl
14	1061.8	94.1	1314	2	US-08-811-949-50	Sequence 50, Appl
15	1061.8	94.1	1974	2	US-08-811-949-38	Sequence 38, Appl
16	1060.4	94.0	1170	2	US-08-811-949-64	Sequence 64, Appl
17	1058.8	93.9	1068	2	US-08-811-949-46	Sequence 46, Appl
18	1057	93.7	1314	2	US-08-811-949-54	Sequence 54, Appl
19	1056.4	93.7	1068	1	US-08-137-116-2	Sequence 2, Appli
20	1054.8	93.5	1065	1	US-08-427-640-1	Sequence 1, Appli
21	1054.8	93.5	1137	4	US-09-553-498-9	Sequence 9, Appli
22	1054.8	93.5	1137	4	US-09-618-869-9	Sequence 9, Appli
23	1054	93.4	1068	2	US-08-811-949-52	Sequence 52, Appl
24	1053.8	93.4	1314	2	US-08-811-949-56	Sequence 56, Appl
25	1052.4	93.3	1170	2	US-08-811-949-66	Sequence 66, Appl
26	1050	93.1	1068	6	523256-3	Patent No. 523256
27	1046.8	92.8	1065	1	US-08-427-640-5	Sequence 5, Appli
28	1038	92.0	1068	2	US-08-811-949-58	Sequence 58, Appl
29	1035.4	91.8	1068	2	US-08-427-640-3	Sequence 3, Appli
30	1030	91.3	1065	2	US-08-811-949-60	Sequence 60, Appl
31	953.4	84.5	2497	6	518259-2	Patent No. 5185259
32	924	81.9	1068	1	US-08-427-640-7	Sequence 7, Appli
33	798	70.7	1163	2	US-08-558-269-5	Sequence 5, Appli
34	798	70.7	1163	3	US-09-410-882-5	Sequence 5, Appli
35	625.8	55.5	1727	6	5244676-4	Patent No. 5244676
36	625.6	55.5	1605	6	5244676-1	Patent No. 5244676
37	472	41.8	472	2	US-08-811-949-40	Sequence 40, Appl
38	414	36.7	453	6	520340-3	Patent No. 5200340
39	267.4	23.7	1724	6	520340-5	Patent No. 5200340
40	170.8	15.1	1236	1	US-07-957-039A-7	Sequence 7, Appli
41	170.8	15.1	1475	4	US-09-643-597-122	Sequence 122, App
42	170.8	15.1	1475	4	US-09-480-884A-122	Sequence 122, App
43	170.8	15.1	1475	4	US-09-542-615A-122	Sequence 122, App
44	170.8	15.1	1475	4	US-09-606-421B-122	Sequence 122, App
45	170.8	15.1	2301	6	5188829-2	Patent No. 5188829

ALIGNMENTS

```
RESULT 1
US-08-811-949-62
; Sequence 62, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;    APPLICANT:  NIWA, MINEO
;    APPLICANT:  SAITO, YOSHIMASA
;    APPLICANT:  SASAKI, HITOSHI
;    APPLICANT:  HAYASHI, MASAKO
;    APPLICANT:  NOTANI, JOUJI
;    APPLICANT:  KOBAYASHI, MASAKAZU
;    TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;    NUMBER OF SEQUENCES: 67
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;       ADDRESSEE: P.C.
;       STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;       CITY: ARLINGTON
;       STATE: VA
;       COUNTRY: USA
;       ZIP: 22202
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/811,949
;       FILING DATE: 05-MAR-1997
;       CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;       NAME: OBLON, NORMAN F.
;       REGISTRATION NUMBER: 24,618
;       REFERENCE/DOCKET NUMBER: 18-966-0
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 703-413-3000
;       TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 62:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 1419 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: circular
;    MOLECULE TYPE: DNA (genomic)
;    FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 1..1416
US-08-811-949-62

  Query Match             94.4%;  Score 1065;  DB 2;  Length 1419;
```

Db 1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1416

RESULT 2
US-08-814-412-10
; Sequence 10, Application US/0881441 2
; Patent No. 6150141
GENERAL INFORMATION:
  APPLICANT: Jarrell Ph.D., Kevin A.
  TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
  NUMBER OF SEQUENCES: 46
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Choate, Hall & Stewart
    STREET: 53 State Street
    CITY: Boston
    STATE: MA
    COUNTRY: USA
    ZIP: 02109
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/814,412
    FILING DATE: 11-MAR-1997
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Jarrell Ph.D., Brenda H.
    REGISTRATION NUMBER: 39,223
    REFERENCE/DOCKET NUMBER: 0079571-0040
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617 248 5000
    TELEFAX: 617 248 4000
INFORMATION FOR SEQ ID NO: 10:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1848 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  IMMEDIATE SOURCE:
    CLONE: amplified t-PA clone
US-08-814-412-10

Query Match 94.4%; Score 1065; DB 3; Length 1848;
Best Local Similarity 99.5%; Pred. No. 8.6e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
         ||| ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  685 CCCCTGCTCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 744

Qy  116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  745 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 804

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  805 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 864

Qy  236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  865 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 924

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCTCCACCTGCGGCGGCTGAGACAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  925 GGCTGACGTGGGAGTACTGTGATGTGCTCCACCTGCGGCGGCTGAGACAGTACA 984

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCAAGAGA 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  985 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCGC 1044
```

Best Local Similarity 99.5%; Pred. No. 7.7e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCCTG 115
         ||| ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCCTG 403

Qy  116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  404 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 463

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  464 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 523

Qy  236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  524 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 583

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCTCCACCTGCGGCGGCTGAGACAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  584 GGCTGACGTGGGAGTACTGTGATGTGCTCCACCTGCGGCGGCTGAGACAGTACA 643

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  644 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGC 703

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGCGGTTCCTGTGCGGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  704 AGGCTGCCATCTTTGCCAAGCACAGGAGCGGTTCCTGTGCGGGGGCA 763

Qy  476 TACTCATCAGCTCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  764 TACTCATCAGCTCCTGCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 823

Qy  536 CCCACCACCTGACGCGTGATCTGGGCAGAACATACCGGGTGCCCTGGGCAGGAGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  824 CCCACCACCTGACGCGTGATCTGGGCAGAACATACCGGGTGCCCTGGGCAGGAGAGC 883

Qy  596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGCACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  884 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGCACTTACGACA 943

Qy  656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCCAGGAGAGCAGCG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  944 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCCAGGAGAGCAGCG 1003

Qy  716 TGGTCCCGCACTGTGTGCCTTCCCCCGGGACCTGCCGGACTGGACGCGGAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1004 TGGTCCCGCACTGTGTGCCTTCCCCCGGGACCTGCCGGACTGGACGCGGAGTGTG 1063

Qy  776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCCGTGAAGG 835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1064 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCCGTGAAGG 1123

Qy  836 AGGCTCATGTCAGACTGTACCCATCCAGCCGTGCACATTCACAACATTTACTTAACAGAA 895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1124 AGGCTCATGTCAGACTGTACCCATCCAGCCGTGCACATTCACAACATTTACTTAACAGAA 1183

Qy  896 CAGTCACCGACAACATGCTGTGTGCTGGGGACACTCGGAGCGGCGGCCCCCAGGCAAACT 955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1184 CAGTCACCGACAACATGCTGTGTGCTGGGGACACTCGGAGCGGCGGCCCCCAGGCAAACT 1243

Qy  956 TGCACGACGGCCTGCCAGGGGATTCGGGGAGGCCCCGGTGTGTCTGAACGATGGCCGCA 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1244 TGCACGACGGCCTGCCAGGGGATTCGGGGAGGCCCCGGTGTGTCTGAACGATGGCCGCA 1303

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1304 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1363

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1710
US-08-811-949-42

Query Match 94.4%;   Score 1065; DB 2; Length 2101;
Best Local Similarity 99.5%;   Pred. No. 9.1e-290;
Matches 1068; Conservative 0; Mismatches 0; Indels 5; Gaps 0;

```
Qy    56 CCCAGGCGCCTCCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
         |||| | || ||||||||||||||||||||||||||||||||||||||||||||||||
Db   638 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 697

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCTCCGTGCCTCCGTGCCTCCGTGAATTCCATGATCCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   698 GCACGCACAGCCTCACCGAGTCGGGTGCCTCTCCGTGCCTCCGTGCCTCCGTGAATTCCATGATCCTGA 757

Qy   176 TAGGCAAGGTTTACACAGCACAGCACCCCAGTGCCCAGGCACTGCCCAGTGCCTGGGCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   758 TAGGCAAGGTTTACACAGCACAGCACCCCAGTGCCCAGGCACTGCCCAGTGCCTGGGCAAACATA 817

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   818 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 877

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGTGCTCACCTGCGGCCTGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   878 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGTGCTCACCTGCGGCCTGAGACAGTACA 937

Qy   356 GCCAGCCTCAGTTTCGCATCAAGGAGGGCTCTTCGCGCCACCCCTCGC 415
         |||||||||||||||||||||||||||||||||||||||||
Db   938 GCCAGCCTCAGTTTCGCATCAAGGAGGGCTCTTCGCGCCAGCACATCGCCCCACCCCTGGC 997

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   998 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1057

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCGCGCCCACTGCTTCCAGGAGGAGAGTTTCCGC 535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1058 TACTCATCAGCTCCTGCTGGATTCTCTCGCGCCCACTGCTTCCAGGAGGAGAGTTTCCGC 1117

Qy   536 CCCACCACCTGACGGTGATCTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1118 CCCACCACCTGACGGTGATCTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 1177

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1178 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1237

Qy   656 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCGGACGCGCG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1238 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCGGACGCGCG 1297

Qy   716 TGGTCCGGACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGACCTGGACGGAGTGTG 775
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1345 TGTCCCGGACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGACCTGGACGGAGTGTG 1404

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCATGAGGGCCTTGCTCCTTTCTATTCGGAGCGCGGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1405 AGTCTCCGGCTACGGCAAGCATGAGGCATGAGGGCCTTGCTCCTTTCTATTCGGAGCGCGGCTGAAGG 1464

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCGCTGCACATCACAACATTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1465 AGGCTCATGTCAGACTGTACCATCCAGCCGCGCTGCACATCACAACATTACTTAACAGAA 1524

Qy   896 CAGTCACCGACAACATGTCGTGTCTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1525 CAGTCACCGACAACATGTCGTGTCTGGAGACACTCGGAGCGGCGGCCCCCAGGCCAAACT 1584

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGGCGCCCCTGTGGTGTCTGAACGATGGCCGCA 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1585 TGCACGACGCCTGCCAGGGCGATTCGGAGGGCGCCCCTGTGGTGTCTGAACGATGGCCGCA 1644

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGTGGGCTGTGGAACAGAGGAGATGTCCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1645 TGACTTTGGTGGGCATCATCAGCTGGGGCGTGGGCTGTGGAACAGAGGAGATGTCCCGGGTG 1704

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1705 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1757
```

RESULT 3
US-08-811-949-42
; Sequence 42, Application US/08811949
; Patent No. 5840533

GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```
Qy    716  TGGTCCGACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1298  TGGTCCGACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  1357

Qy    776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG  835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1358  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG  1417

Qy    836  AGGCTCATGTCAGACTGTACCCATCCAGCGCTGCACATCACAACATTTACTTAACAGAA  895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1418  AGGCTCATGTCAGACTGTACCCATCCAGCGCTGCACATCACAACATTTACTTAACAGAA  1477

Qy    896  CAGTCACCGACAACAATGCTGTGTGCTCGGAGACACTCGGAGCGCGGCGCCCAGGCAAACT  955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1478  CAGTCACCGACAACAATGCTGTGTGCTCGGAGACACTCGGAGCGCGGCGCCCAGGCAAACT  1537

Qy    956  TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTGTCTGAACGATGGCCGCA  1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1538  TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTGTCTGAACGATGGCCGCA  1597

Qy   1016  TGACTTTGGTGGGCATCATCAGCTGGGGCCTCGGGCCTCGGGGCCTCGGGCGCCCGGGTG  1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1598  TGACTTTGGTGGGCATCATCAGCTGGGGCCTCGGGCCTCGGGGCCTCGGGCGCCCGGGTG  1657

Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1710


RESULT 4
US-08-119-512-3
; Sequence 3, Application US/08119512
; Patent No. 5498531
GENERAL INFORMATION:
;   APPLICANT:  Jarrell, Kevin A.
;   TITLE OF INVENTION:  INTRON MEDIATED RECOMBINANT TECHNIQUES
;   TITLE OF INVENTION:  AND REAGENTS
;   NUMBER OF SEQUENCES:  3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  LAHIVE & COCKFIELD
;   STREET:  60 State Street
;   CITY:  Boston
;   STATE:  MA
;   COUNTRY:  USA
;   ZIP:  02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  ASCII (text)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/119,512
;   FILING DATE:  10-SEP-1993
;   CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Vincent, Matthew P.
;   REGISTRATION NUMBER:  36,709
;   REFERENCE/DOCKET NUMBER:  HUI-008
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (617) 227-7400
;   TELEFAX:  (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  2162 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  double
;   TOPOLOGY:  both
;   MOLECULE TYPE:  other nucleic acid
;   FEATURE:
;   NAME/KEY:  misc feature
;   LOCATION:  82..334
;   OTHER INFORMATION:  /product= "Signal Sequence and
;   OTHER INFORMATION:  Finger-like domain"
```

```
;   FEATURE:
;   NAME/KEY:  misc feature
;   LOCATION:  335..447
;   OTHER INFORMATION:  /product= "EGF-like domain"
;   FEATURE:
;   NAME/KEY:  misc feature
;   LOCATION:  448..714
;   OTHER INFORMATION:  /product= "Kringle-1 domain"
;   FEATURE:
;   NAME/KEY:  misc feature
;   LOCATION:  715..972
;   OTHER INFORMATION:  /product= "Kringle-2 domain"
;   FEATURE:
;   NAME/KEY:  misc feature
;   LOCATION:  973..2162
;   OTHER INFORMATION:  /product= "Catalytic domain"
US-08-119-512-3

Query Match            94.4%;    Score 1065;  DB 1;  Length 2162;
Best Local Similarity  99.5%;    Pred. No. 9.2e-290;
Matches 1068;  Conservative   0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy     56  CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTG  115
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    695  CCCCTGCTGTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  754

Qy    116  GCACGCACAGCCTCACCGAGTCGGGTGCGTCGGGTGCCTCCTGCCTTCCCGTGCCTCCCGTGCTGAATTCCATGATCCTGA  175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    755  GCACGCACAGCCTCACCGAGTCGGGTGCGTCGGGTGCCTCCTGCCTTCCCGTGCCTCCCGTGCTGAATTCCATGATCCTGA  814

Qy    176  TAGGCAAGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGCCGGCCTGGCCACTGCC  235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    815  TAGGCAAGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGCCGGCCTGGCCACTGCC  874

Qy    236  ATTACTGCGGAATCCTGATGGGCAGCCTGGTGCCACGTGCTGACGAACCGCA  295
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    875  ATTACTGCGGAATCCTGATGGGCAGCCTGGTGCCACGTGCTGAAGAACCGCA  934

Qy    296  GGCTGACGTGGGAGTACTGTGATGTGCCTCCTGCTCCACCTGCCGGTGCTGAGACAGTACA  355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    935  GGCTGACGTGGGAGTACTGTGATGTGCCTCCTGCTCCACCTGCCGGTGCTGAGACAGTACA  994

Qy    356  GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    995  GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  1054

Qy    416  AGGCTGCCATCTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCGTGTGGGGGCA  475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055  AGGCTGCCATCTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCGTGTGGGGGCA  1114

Qy    476  TACTCATCAGCTCCTGCGGATTCTCTCGCCGCCACTGCTTCTCCAGGAGAGGTTTCGC  535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115  TACTCATCAGCTCCTGCGGATTCTCTCGCCGCCACTGCTTCTCCAGGAGAGGTTTCGC  1174

Qy    536  CCCACCACCTGACGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC  595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175  CCCACCACCTGACGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC  1234

Qy    596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1294

Qy    656  ATGACATTGCGCTGCAGCTGAAATCGGATTCGTCCGGTCGTGTGCCAGGAGCAGCG  715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295  ATGACATTGCGCTGCAGCTGAAATCGGATTCGTCCGGTCGTGTGCCAGGAGCAGCG  1354

Qy    716  TGGTCCGACTGTGTGCCTTCCCCGGGGACTGCCAGCTGCCGGACTGGACGGAGTGTG  775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355  TGGTCCGACTGTGTGCCTTCCCCGGGGACTGCCAGCTGCCGGACTGGACGGAGTGTG  1414

Qy    776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG  835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCTTTCTATTCGGAGCCGCTGAAGG  1474
```

```
        OTHER INFORMATION:  /product= "Kringle-1 domain"
      FEATURE:
        NAME/KEY: misc_feature
        LOCATION: 715..972
        OTHER INFORMATION:  /product= "Kringle-2 domain"
      FEATURE:
        NAME/KEY: misc_feature
        LOCATION: 973..2162
        OTHER INFORMATION:  /product= "Catalytic domain"
US-08-488-015B-3

Query Match           94.4%;  Score 1065;  DB 1;  Length 2162;
Best Local Similarity 99.5%;  Pred. No. 9.2e-290;
Matches 1068;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

Qy    56 CCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  115
            ||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 CCCCTGCCTGTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  754

Qy   116 GCACGCAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGA  175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755 GCACGCAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGCCTCCCGTGAATTCCATGATCCTGA  814

Qy   176 TAGGCAAGTTTACACAGCACAGACCCCAGTGCCAGCACTGGGCCTGGGCAAACATA  235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815 TAGGCAAGTTTACACAGCACAGACCCCAGTGCCAGCACTGGGCCTGGGCAAACATA  874

Qy   236 ATTACTGCCGGAATCCTGATGGGATGGGATGTGCCCCTGCCACGTGCTGAAGAACCGCA  295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   875 ATTACTGCCGGAATCCTGATGGGATGGGATGTGCCCCTGCCACGTGCTGAAGAACCGCA  934

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTGCTCCAACTGCCGCCTGAGACAGTACA  355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   935 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTGCTCCAACTGCCGCCTGAGACAGTACA  994

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC  415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 1054

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCGTGTCGCGGGGCA  475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1055 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCGTGTCGCGGGGCA 1114

Qy   476 TACTCATCAGCTCCGTGCTGGATTCTCTGCGCCAGCCACTGCTTCCAGGAGAGTTTCCGC  535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1115 TACTCATCAGCTCCGTGCTGGATTCTCTGCGCCAGCCACTGCTTCCAGGAGAGTTTCCGC 1174

Qy   536 CCCACCACTGACGGTGATCTTCGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC  595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1175 CCCACCACTGACGGTGATCTTCGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1234

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGACGAGCG  715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1295 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGACGAGCG 1354

Qy   716 TGGTTCCGGACACTGTGTGCCTTCCCCGGGGACCTGCAGCTGCCAGCTGGACCGGAGTGTG  775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1355 TGGTTCCGGACACTGTGTGCCTTCCCCGGGGACCTGCAGCTGCCAGCTGGACCGGAGTGTG 1414

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1415 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1474

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA  895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1475 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1534

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCGCCCCAGGCAAACT  955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1535 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCGCCCCAGGCAAACT 1594

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCGTGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1595 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCGTGTGTCTGAACGATGGCCGCA 1654

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGGCCTGGGCCTGGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1655 TGACTTTGGTGGGCATCATCAGCTGGGGGCCTGGGCCTGGACAGAAGGATGTCCCGGGTG 1714

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGCGACCG 1767
```

Db   875 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGCCACGTGCTGAGAGAACCGCA 934
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1535 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGACGCGGCGGCCCCAGGCAAACT 1594

Qy   296 GGGTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTGTCTCGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1015

Db   935 GGGTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTGTCTCCACCTGCCTGAGACAGTACA 994
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1595 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1654

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCCGACATCGCCTCCACCCCTGGC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1016 TGACTTTGGTGGGCATCATCATCAGCTGGGCTGTGGGCGACAGAGGATGTCCGGGTG 1075

Db   995 GCCAGCCTCAGTTTCGCATCAAAGAGGAGGCTCTTCGCCCGACATCGCCTCCACCCCTGGC 1054
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1655 TGACTTTGGTGGGCATCATCATCAGCTGGGCTGTGGGCGACAGAGGATGTCCGGGTG 1714

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128

Db 1055 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA 1114
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1767

Qy   476 TACTCATCAGCTCCTGCTGGAATTCTCTGCTGGAATTCTCTGCTGTCTGCTTTCTCCGCT 535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1115 TACTCATCAGCTCCTGCTGGAATTCTCTGCTGGAATTCTCTGCTGTCTGCTTTCCGC 1174

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1175 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1234

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCTGCTGCCAGGAGCAGCGG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1295 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCTGCTGCCAGGAGCAGCGG 1354

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGCCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1355 TGGTCCGCACTGTGTGCCTTCCCCGGCCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1414

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGGCTGAAGG 835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1415 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGGCTGAAGG 1474

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1475 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1534

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACTCGGAGGCCCCCAGGCAAACT 955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1535 CAGTCACCGACAACATGCTGTGTGCTGGAGACTCGGAGGCCCCCAGGCAAACT 1594

Qy   956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCGATTCGGGAGGCCGATGTGTCTGAACGATGCCGCA 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1595 TGCACGACGCCTGCCAGGCGATTCGGGAGGCGATTCGGGAGGCCGATGTGTCTGAACGATGCCGCA 1654

Qy 1016 TGACTTTGGTGGGCATCATCATCAGCTGGGCTGTGGGGCTGTGTGACAGAAGGATGTCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1655 TGACTTTGGTGGGCATCATCATCAGCTGGGCTGTGGGGCTGTGTGACAGAAGGATGTCCGGGTG 1714

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1767

RESULT 7
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:

RESULT 6
US-08-488-015B-25
; Sequence 25, Application US/08488015B
; Patent No. 5780212
; GENERAL INFORMATION:
;   APPLICANT: Jarrell, Kevin A.
;   TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
;   TITLE OF INVENTION: AND REAGENTS
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley, Hoag & Eliot
;   STREET: One Post Office Square
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII (text)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/488,015B
;   FILING DATE: 07-JUN-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Vincent, Matthew P.
;   REGISTRATION NUMBER: 36,709
;   REFERENCE/DOCKET NUMBER: HUV-008.02
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617) 832-1000
;   TELEFAX: (617) 832-7000
;   INFORMATION FOR SEQ ID NO: 25:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2162 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-488-015B-25

Query Match   94.4%; Score 1065; DB 1; Length 2162;
Best Local Similarity 99.5%; Pred. No. 9.2e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTGGTCAGCCTACCGTG 115
         ||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTGGTCAGCCTACCGTG 754

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGCCTCCGTGATCCTGA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755 GCACGCACCAGCCTCACCGAGTCGGGTGCGGTGCCTCCTGCCTCCGTGCCTCCGTGATCCTGA 814

Qy   176 TAGGCAAGGTTACACAGCCAGACAGACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815 TAGGCAAGGTTACACAGCCAGACAGACCCCAGTGCCCAGCACTGGGCCTGGGCAAACATA 874

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGCCACGTGCTGAGAAGAACCCA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
;       ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;       ADDRESSEE:  P.C.
;       STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;       CITY:  ARLINGTON
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22202
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/811,949
;       FILING DATE:  05-MAR-1997
;       CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;       NAME:  OBLON, NORMAN F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  18-966-0
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-413-3000
;       TELEFAX:  703-413-2220
; INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  1314 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1311
US-08-811-949-48

Query Match            94.3%;  Score 1063.4;  DB 2;  Length 1314;
Best Local Similarity  99.4%;  Pred. No. 2.1e-289;
Matches 1067;  Conservative  0;  Mismatches  6;  Indels  0;  Gaps  0;

Qy      56 CCCAGGCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGCTCAGCCTACCGTG 115
           |||   ||  ||| ||||||||||||||||||| ||||||||||||| |||||||||||||
Db     239 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGGTCAGCCTACCGTG 298

Qy     116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 358

Qy     176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     359 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 418

Qy     236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCCAGTGCTGAAG... 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     419 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCCAGTGCTGAAG 478

Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCGCGCCTCCACCTGCGCGGCCTGAGACAGTACA 355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479 GGCTGACGTGGGAGTACTGTGATGTGCGCGCCTCCACCTGCGCGGCCTGAGACAGTACA 538

Qy     356 GCCAGGCTCAGTTTCGCATCAAAGGAGGGCTTCGCCGCACATCGCCCCACCCCTGGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     539 GCCAGGCTCAGTTTCGCATCAAAGGAGGGCTTCGCCGCACATCGCCCCACCCCTGGC 598

Qy     416 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGAGCGGTTCCTGTGCGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     599 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGAGCGGTTCCTGTGCGGGGGCA 658

Qy     476 TACTCATCAGCTCCTGCTGGAATTCTCTCTGCCGCCGCCACTGCTTCTGCAGAGAGTTTCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     659 TACTCATCAGCTCCTGCTGGAATTCTCTCTGCCGCCGCCACTGCTTCTGCAGAGAGTTTCCGC 718

Qy     536 CCCACCACCTGACGCGTGATCTTGGGCAGGAACATACCGGTGTGTCCCTGGCGGAGGAGAGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 8
5200340-1
;Patent No. 5200340
;  APPLICANT:  FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
;  TITLE OF INVENTION:  THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;  NUMBER OF SEQUENCES:  34
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/07/53,412
;     FILING DATE:  22-MAY-1987
;SEQ ID NO:1:
;  LENGTH:  1738
5200340-1

Query Match            94.3%;  Score 1063.4;  DB 6;  Length 1738;
Best Local Similarity  99.4%;  Pred. No. 2.4e-289;
Matches 1067;  Conservative  0;  Mismatches  6;  Indels  0;  Gaps  0;

Qy      56 CCCAGGCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           |||   ||  ||| |||||||||||||||| |||||||||||||||||||||||||||||
Db     629 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGGTCAGCCTACCGTG 688

Qy     116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     689 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 748

Qy     176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     749 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 808

Qy     236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGACCGCA 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     809 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGACCGCA 868

Db     719 CCCACCACCTGACGCGTGATCTTGGGCAGCAACATACCGGTGTGTCCCTGCCGGAGGAGAGC 778
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           .||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     779 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 838

Qy     656 ATGACAATTGCGCTGCGACTGTGTCCGATTCGGATTCGTCCCGCTGTGCCCGAGGAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     839 ATGACAATTGCGCTGCGACTGTGTCCGATTCGGATTCGTCCCGCTGTGCCCGAGGAGCAGCG 898

Qy     716 TGGTCCGCACTGTGTCCTTCCCCGGCGGCGGACTGCGACTGCGACGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     899 TGGTCCGCACTGTGTCCTTCCCCGGCGGCGGACTGCGACTGCGACGGAGTGTG 958

Qy     776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     959 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG 1018

Qy     836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1019 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1078

Qy     896 CAGTCACCGACACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCCAGGCAAACT 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1079 CAGTCACCGACACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCCAGGCAAACT 1138

Qy     956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1139 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1198

Qy    1016 TGACTTTGGTGGCATCATCAGCTGGGCCTGGGGCGTGTTGGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1199 TCACTTTGGTGGCATCATCAGCTGGGCCTGGGGCGTGTTGGACAGAAGGATGTCCCGGGTG 1258

Qy    1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1259 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1311
```

```
                COUNTRY:  Canada
                ZIP:  M5H 3Y2
        COMPUTER READABLE FORM:
                MEDIUM TYPE:  Floppy disk
                COMPUTER:  IBM PC compatible
                OPERATING SYSTEM:  PC-DOS/MS-DOS
                SOFTWARE:  PatentIn Release #1.0, Version #1.25
        CURRENT APPLICATION DATA:
                APPLICATION NUMBER:  US/08/883,795A
                FILING DATE:  27-JUN-1997
                CLASSIFICATION:  435
        ATTORNEY/AGENT INFORMATION:
                NAME:  Gravelle, Micheline
                REGISTRATION NUMBER:  40,261
                REFERENCE/DOCKET NUMBER:  7841-062
        TELECOMMUNICATION INFORMATION:
                TELEPHONE:  (416) 364-7311
                TELEFAX:  (416) 361-1398
        INFORMATION FOR SEQ ID NO:  39:
        SEQUENCE CHARACTERISTICS:
                LENGTH:  1955 base pairs
                TYPE:  nucleic acid
                STRANDEDNESS:  single
                TOPOLOGY:  linear
        MOLECULE TYPE:  cDNA
        ORIGINAL SOURCE:
                ORGANISM:  Homo sapiens
     ; US-08-883-795A-39

Query Match             94.3%;  Score 1063.4;  DB 2;  Length 1955;
Best Local Similarity   99.4%;  Pred. No. 2.5e-289;
Matches 1067;  Conservative  0;  Mismatches  0;  Indels  6;  Gaps  0;
```

```
Qy      56 CCCAGGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     704 CCCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 763

Qy     116 GCACGCACAGCCTCACCGAGTGGGGTGCCTCCTGCCTCCCGTGGCCTCCGTGGAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     764 GCACGCACAGCCTCACCGAGTGGGGTGCCTCCTGCCTCCCGTGGCCTCCGTGGAATTCCATGATCCTGA 823

Qy     176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     824 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 883

Qy     236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     884 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 943

Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     944 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 1003

Qy     356 GCCAGCCTCAGTTTCGCATCAAGAGCACAGGAGTCGCCGGAGAGCCGGTTCCGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1004 GCCAGCCTCAGTTTCGCATCAAGAGCACAGGAGTCGCCGGAGAGCCGGTTCCGC 1063

Qy     416 AGGCTGCCATCTTGCCAAGCACAGGACGCGTTCCTGTCGCGGGGCA 475
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1064 AGGCTGCCATCTTGCCAAGCACAGGACGCGTTCCTGTCGCGGGGCA 1123

Qy     476 TACTCATCAGCTCCTGCGGATTCTCTCTGCGCCGCCCACTGCTTCTCTGCGCC 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1124 TACTCATCAGCTCCTGCGGATTCTCTCTGCGCCGCCCACTGCTTCTCTGCGCC 1183

Qy     536 CCCACCACCTGACCGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGCCGGAGGAGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1184 CCCACCACCTGACCGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGCCGGAGGAGC 1243

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1244 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1303
```

```
Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     869 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 928

Qy     356 GCCAGCCTCAGTTTCGCATCAAAGGAGGCTCTTCGCGCGACATGCCCCACCCCTGGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     929 GCCAGCCTCAGTTTCGCATCAAAGGAGGCTCTTCGCGCGACATGCCCCACCCCTGGC 988

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGGCCCGGAGAGCACGGTTCCTGTGCGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     989 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGGCCCGGAGAGCACGGTTCCTGTGCGGGGCA 1048

Qy     476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCTCTCTGCGCCCACTGCTTCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1049 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCTCTCTGCGCCCACTGCTTCCGC 1108

Qy     536 CCCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGTCCCTGGCGAGGAGC 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1109 CCCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGTCCCTGGCGAGGAGC 1168

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1169 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1228

Qy     656 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1229 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1288

Qy     716 TGGTCCGACTGTGTGCCTTCCCCGGGACCTGCCGGACTGGACGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1289 TGGTCCGACTGTGTGCCTTCCCCGGGACCTGCCGGACTGGACGGAGTGTG 1348

Qy     776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGCTCCTTCCTTCATTCGGAGCGCGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1349 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGCTCCTTCCTTCATTCGGAGCGCGCTGAAGG 1408

Qy     836 AGGCTCATGTACCCAGCGCGTGCACATCACAACATTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    1409 AGGCTCATGTACCCAGCGCGTGCACATCACAACATTACTTAACAGAA 1468

Qy     896 CAGTCACCGACAACACATGCTGTGCTGAGACACTCGGAGCGGCGGCCCAGGCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1469 CAGTCACCGACAACACATGCTGTGCTGAGACACTCGGAGCGGCGGCCCAGGCAAACT 1528

Qy     956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGGATTCGGGAGGCGGATTCGGGAGGCGGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1529 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGGATTCGGGAGGCGGATTCGGGAGGCGGGCCGCA 1588

Qy    1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGGGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1589 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGGGACAGAAGGATGTCCCGGGTG 1648

Qy    1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1649 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1701
```

```
RESULT 9
US-08-883-795A-39
; Sequence 39, Application US/08883795A
; Patent No. 5985607
GENERAL INFORMATION:
        APPLICANT:  Delcuve, Genevieve
        APPLICANT:  Awang, Gregor
        TITLE OF INVENTION:  Recombinant DNA Molecules and Expression
        TITLE OF INVENTION:  Vectors for Tissue Plasminogen Activator
        NUMBER OF SEQUENCES:  39
        CORRESPONDENCE ADDRESS:
                ADDRESSEE:  BERESKIN & PARR
                STREET:  40 King Street West
                CITY:  Toronto
                STATE:  Ontario
```

```
Qy  656 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCAGGAGAGCAGCG  715
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1304 ATGACAATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCAGGAGAGCAGCG 1363

Qy  716 TGGTCCGCACTGTGTGCCTTCGCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1364 TGGTCCGCACTGTGTGCCTTCGCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1423

Qy  776 AGCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCTTTCTATTCGGAGCGGCTGAAGG  835
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1424 AGCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCTTTCTATTCGGAGCGGCTGAAGG 1483

Qy  836 AGGCTCATGTCAGACTGTACCATCCAGCCGCCGCTGCACATCACAAACATTTACTTAACAGAA  895
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1484 AGGCTCATGTCAGACTGTACCATCCAGCCGCCGCTGCACATCACAAACATTTACTTAACAGAA 1543

Qy  896 CAGTCACCGACAACATGCTGTGTCGAGACACTCGGAGCCGGCGGACCCCAGGCCAAACT  955
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1544 CAGTCACCGACAACATGCTGTGTCGAGACACTCGGAGCCGGCGGACCCCAGGCCAAACT 1603

Qy  956 TGCACGACGCCTGCCAGGCGGATTCGGGAGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1015
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1604 TGCACGACGCCTGCCAGGCGGATTCGGGAGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1663

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1664 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1723

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1724 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1776
```

RESULT 10
5344773-1
;Patent No. 5344773
;APPLICANT: WEI, CHA-MER;HSIUNG, NANCY;REDDY, VERMURI B.;
;LEMONTT, JEFFREY F.;DACKOWSKI, WILLIAM;DOUGLAS, RICHARD;
;COLE, EDWARD S.;PURCELL JR., RICHARD D.;LAU, DAVID TAI-YUI
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
;ACTIVATOR PRODUCED BY RECOMBINANT DNA
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/782,686
; FILING DATE: 01-OCT-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656,770
; FILING DATE: 01-OCT-1984
;SEQ ID NO:1:
;LENGTH: 2457
5344773-1

Query Match 94.3%; Score 1063.4; DB 6; Length 2457;
Best Local Similarity 99.4%; Pred. No. 2.7e-289;
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy   56 CCCAGGCGGCCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
        |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  627 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  686

Qy  116 GCACGCCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA  175
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  687 GCACGCCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA  746

Qy  176 TAGGCAAGGTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCAAACATA  235
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  747 TAGGCAAGGTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCCAAACATA  806

Qy  236 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTTGGTGCCACGTGCTGAAGAACCGCA  295
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  807 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTTGGTGCCACGTGCTGAAGAACCGCA  866

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCCCTCCACCTGCGGCCTGAGACAGTACA  355
```

RESULT 11
US-08-286-740-1
; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
; APPLICANT: Crowley, Craig W.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:

```
Qy  867 GGCTGACGTGGGAGTACTGTGATGTGCCCATCAACCTGCCCTCCTGCTCCACCTGCTGGACAGTACA  926
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  356 GCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCGACATCGCTCTCGCTCCCACCCCTGGC  415

Qy  927 GCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  986
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  416 AGGCTGCCATCTTCCAAGCACAGGAGTCGCCGGAGAGCGGTTCGTGTGCGGGGGCA  475

Qy  416 AGGCTGCCATCTTCCAAGCACAGGAGTCGCCGGAGAGCGGTTCGTGTGCGGGGGCA  475
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  987 AGGCTGCCATCTTCCAAGCACAGGAGTCGCCGGAGAGCGGTTCGTGTGCGGGGGCA 1046

Qy  476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCAGCACAACATTTACTTAACAGAA  535
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1047 TACTCATCAGCTCCTGCTGGATTCTCTCTGCAGCACAACATTTACTTAACAGAA 1106

Qy  536 CCCACCACCTGACGGTGATCTTGGCAGAACATACCGGGTGCCCTGGCGGGGAGGAGC  595
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1107 CCCACCACCTGACGGTGATCTTGGCAGAACATACCGGGTGCCCTGGCGGGGAGGAGC 1166

Qy  596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1167 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1226

Qy  656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCAGGAGAGCAGCG  715
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1227 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCAGGAGAGCAGCG 1286

Qy  716 TGGTCCGCACTGTGTGCCTTCGCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1287 TGGTCCGCACTGTGTGCCTTCGCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1346

Qy  776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1347 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1406

Qy  836 AGGCTCATGTCAGACTGTACCATCCAGCCGCCGCTGCACATCACAACATTTACTTAACAGAA  895
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1407 AGGCTCATGTCAGACTGTACCATCCAGCCGCCGCTGCACATCACAACATTTACTTAACAGAA 1466

Qy  896 CAGTCACCGACAACATGCTGTGTCGTGTCGAGACACTCGGAGCGGCGGCCCCAGGCAAACT  955
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1467 CAGTCACCGACAACATGCTGTGTCGTGTCGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1526

Qy  956 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCTGTGTGTCTGAACGATGGCCGCA 1015
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1527 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCTGTGTGTCTGAACGATGGCCGCA 1586

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCGTGTGGACGAAGGATGTCCCGGGTG 1075
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1587 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCGTGTGGACGAAGGATGTCCCGGGTG 1646

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTG 1128
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1647 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACACCGTG 1699
```

;   MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/286,740
;   FILING DATE: 05-AUG-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Lee, Wendy M.
;   REGISTRATION NUMBER:
;   REFERENCE/DOCKET NUMBER: 798
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-1994
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 7360 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
US-08-286-740-1

Query Match   94.3%;   Score 1063.4;   DB 1;   Length 7360;
Best Local Similarity   99.4%;   Pred. No. 4.3e-289;
Matches 1067;   Conservative   0;   Mismatches   0;   Indels   6;   Gaps   0;

Qy   56 CCCAGGGGGCCTCTGAGGGAAACAGTGACTGTACTTTGGGAATGGTCAGCCTACCGTG 115
          ||||| |||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   2235 CCCCTGCCTGCTGAGGGAAACAGTGACTGCTGCTACTTTGGGAATGGGTCAGCCTACCGTG 2294

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCCGTGGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2295 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCCGTGGAATTCCATGATCCTGA 2354

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2355 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 2414

Qy   236 ATTACTGCCGGAATCCTGATGGGGGATGGGCCAAGCCCTGGTGCCACGTGCTGCGGCCCA 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2415 ATTACTGCCGGAATCCTGATGGGGGATGGGCCAAGCCCTGGTGCCACGTGCTGCGGCCCA 2474

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCCACCTGCGGCTCCGAGACAGTACA 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2475 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCCACCTGCGGCTCCGAGACAGTACA 2534

Qy   356 GCCAGCCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2535 GCCAGCCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 2594

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGTGGAGAGCGGTTCCTGTGCGGGGGCA 475
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2595 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGCCGGAGAGCGGTTCCTGTGCGGGGGCA 2654

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCCACTGCTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2655 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCCACTGCTTCCGC 2714

Qy   536 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2715 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 2774

Qy   596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2775 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 2834

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2835 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 2894

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGGACCTGCAGCTGCGACTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2895 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGGACCTGCAGCTGCGACTGGACGGAGTGTG 2954

Qy   776 AGCTCTCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2955 AGCTCTCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCGGCTGAAGG 3014

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3015 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 3074

Qy   896 CAGTCACCACGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGCCCAGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3075 CAGTCACCACGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGCCCAGCAAACT 3134

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTCGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3135 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTCGTGTGTCTGAACGATGGCCGCA 3194

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGGCTGGGCGTGGGACAGAAGGATGTCCCGGGGTG 1075
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3195 TGACTTTTGGTGGGCATCATCAGCTGGGGCGGCTGGGCGTGGGACAGAAGGATGTCCCGGGGTG 3254

Qy   1076 TGTACACAAAGGTTACCAACTACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3255 TGTACACCAAGGTTACCAACTACTACCTAGACTGGATTCGTGACAACATGCCGACCG 3307

RESULT 12
PCT-US95-09576-1
; Sequence 1, Application PC/TUS9509576
; GENERAL INFORMATION:
;   APPLICANT: GENENTECH, INC.
;   TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
;   TITLE OF INVENTION: HOST CELLS
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US95/09576
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/286740
;   FILING DATE: 05-AUG-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Lee, Wendy M.
;   REGISTRATION NUMBER: 00,000
;   REFERENCE/DOCKET NUMBER: 798PCT
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-1994
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 7360 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
PCT-US95-09576-1

```
Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   3255  TGTACACCAAGGTTACCAACTACCTAGACTACCTAGACTGGATTCGTGACAACATGCGACCG  3307
```

RESULT 13
US-08-811-949-44
; Sequence 44, Application US/0811949
; Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 44:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1065
US-08-811-949-44

Query Match   94.1%; Score 1062; DB 2; Length 1068;
Best Local Similarity 100.0%; Pred. No. 4.8e-289;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     67  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGC   126
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4  TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGC    63

Qy    127  CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   186
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64  CTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGATAGGCAAGGTT   123

Qy    187  TACACAGCACAGAACCCCAGTGGCACTGGGCCTCGGCAAACATTACTTACTGCCCG     246
           ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    124  TACACAGCACAGAACCCCAGTGGCACTGGGCCTCGGCAAACATATATTACTGCCCG     183

Qy    247  AATCCTGATGGGGATGCCAAGCCCTGGTGCCCACGTGCTGAAGAACCGCAGGCTGACGTGG   306
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184  AATCCTGATGGGGATGCCAAGCCCTGGTGCCCACGTGCTGAAGAACCGCAGGCTGACGTGG   243
```

Query Match   94.3%; Score 1063.4; DB 5; Length 7360;
Best Local Similarity 99.4%; Pred. No. 4.3e-289;
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy     56  CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTG   115
           ||||| ||  |||||||||||||||||||||||||||||||||||||||||||||||
Db   2235  CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTG  2294

Qy    116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGA   175
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2295  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGA  2354

Qy    176  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGGCACTGGGCCTGGGCAAACATA   235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2355  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGGCACTGGGCCTGGGCAAACATA  2414

Qy    236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCCACGTGCTGAAGAACCGCA   295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2415  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCCACGTGCTGAAGAACCCCA  2474

Qy    296  GGCTGACGTGGGGAGTACTGTGATGTGCCCTCGTGCTCCACCTGCGCGGCCTGAGACAGTACA   355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2475  GGCTGACGTGGGGAGTACTGTGATGTGCCCTCGTGCTCCACCTGCGCGGCCTGAGACAGTACA  2534

Qy    356  GCCAGCCTCAGTTTCGCATCAAAGAGAGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC   415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2535  GCCAGCCTCAGTTTCGCATCAAAGAGAGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  2594

Qy    416  AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGTTCCTGTGGGGGGCA   475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2595  AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCA  2654

Qy    476  TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGC   535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2655  TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCCCCACTGCTTCCAGGAGAGGTTTCCGC  2714

Qy    536  CCCCACCACTGACGGTGATCTTTGGGCAGCAGAACATACCGGGTCCTGGCGTCCTGGAGGAGC   595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2715  CCCCACCACTGACGGTGATCTTTGGGCAGCAGAACATACCGGGTCCTGGCGTGGTCCTGGAGGAGC  2774

Qy    596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA   655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2775  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  2834

Qy    656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTTCCCGCGTGTGCCCAGGAGAGCAGCG   715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2835  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTTCCCGCGTGTGCCCAGGAGAGCAGCG  2894

Qy    716  TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG   775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2895  TGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  2954

Qy    776  AGTCTCCGGCTACGGCAAGCATGACAAGCATGAGGCCTTGTCTCTTGTCTCCTTCTATTCCGAGGCCGCTGAAGG   835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2955  AGTCTCCGGCTACGGCAAGCATGACAAGCATGAGGCCTTGTCTCTTGTCTCCTTCTATTCCGAGGCCGCTGAAGG  3014

Qy    836  AGGCTCATGTCAGACTGTCAGACTTACTTTACTTTACTTTAACAGAA   895
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   3015  AGGCTCATGTCAGACTGTCAGACTTACTTTACTTTACTTTAACAGAA  3074

Qy    896  CAGTCACCGACAACAACATGCTGTGTGCTGAGACACTCGGAGCCGCCGGCCCCAGGCAAACT   955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3075  CAGTCACCGACAACAACATGCTGTGTGCTGAGACACTCGGAGCCGCCGGCCCCAGGCAAACT  3134

Qy    956  TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGGCCGCA   1015
           |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   3135  TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTTGGTGTGTCTGAACGATGGGCCGCA  3194

Qy   1016  TGACTTTTGGTGGGCATCATCAGCTGGGGCTGGGCCTGGGGCTGTGGACAGAAGGATGTCCCGGGTG   1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3195  TGACTTTTGGTGGGCATCATCAGCTGGGGCTGGGCCTGGGGCTGTGGACAGAAGGATGTCCCGGGTG  3254
```

```
  ADDRESSEE: P.C.
  STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
  CITY: ARLINGTON
  STATE: VA
  COUNTRY: USA
  ZIP: 22202
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/811,949
  FILING DATE: 05-MAR-1997
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: OBLON, NORMAN F.
  REGISTRATION NUMBER: 24,618
  REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 703-413-3000
  TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 50:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1314 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: circular
  MOLECULE TYPE: DNA (genomic)
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 1..1311
US-08-811-949-50

Query Match             94.1%;  Score 1061.8;  DB 2;  Length 1314;
Best Local Similarity   99.3%;  Pred. No. 6e-289;
Matches 1066;  Conservative 0;  Mismatches 7;  Indels 0;  Gaps 0;

Qy    56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGACTTGGGAATGGTCAGCCTACCGTG  115
            ||||| ||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   239 CCCCTGCTGCTCTGAGGGAAACAGTGACTGACTGCTACTTGGGAATGGTCAGCCTACCGTG  298

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCGTGCCTCCGTGCCTCCGTGAATTCCATGATCCTGA  175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GCACGCACAGCCTCACCGAGTCGGGTGCGTGCCTCCGTGCCTCCGTGAATTCCATGATCCTGA  358

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  418

Qy   236 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCACGTGCCAAGACCCGCA  295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCACGTGCCAAGACCCGCA  478

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCCTCGTCCACCTGGGCCTGACCTGAGACAGTAGA  355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 GGCTGACGTGGGAGTACTGTGATGTGCCCCTCGTCCACCTGGGCCTGACCTGAGACAGTAGA  538

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGCACATGCCACCCCTGGC  415
          |||||||||||||||||||||| ||||||| |||||||||||||||||||||||||
Db   539 GCCAGCCTCAGTTTCGCATCATAGGAGGCCTCTCCGCCGCACATGCCACCCCTGGC  598

Qy   416 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGTTCCTGTGCGGGGGCA  475
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGTTCCTGTGCGGGGGCA  658

Qy   476 TACTCATCAGCTCCTGCTTCTCTGCGCCCCACTGCTTCCAGGAGAGTTTCCGC  535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TACTCATCAGCTCCTGCTTCTCTGCGCCCCACTGCTTCCAGGAGAGTTTCCGC  718

Qy   536 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGGGAGC  595
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   719 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC  778

Qy   307 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCTGAGACAGTACAGCCAGCCTCAG  366
          ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db   244 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCCGGCCTGAGACAGTACAGCCAGCCTCAG  303

Qy   367 TTTCGCATCAAAGGAGGAGGCTCTTCGCCGCACATGCCGCACCTGCCATC  426
          |||||||||||||||||||| |||||||||||||||||||||||||||||
Db   304 TTTCGCATCAAAGGAGGAGGGCTCTTCGCCGCACATGCCGCACCTGCCATC  363

Qy   427 TTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC  486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTCATCAGC  423

Qy   487 TCCTGCTGGATTCTCTGCCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG  546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 TCCTGCTGGATTCTCTGCCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG  483

Qy   547 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCTGGAGGAGCAGAAATTTGAA  606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCTGGAGGAGCAGAAATTTGAA  543

Qy   607 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTCGACAATGACATTGCG  666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTCGACAATGACATTGCG  603

Qy   667 CTGCTGCAGCTGAAATCGGATTGTCCCGGTGCCGTGGTCAGCGTGGTCCGCGACT  726
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   604 CTGCTGCAGCTGAAATCGGATTGTCCCGTGCCGTGGTCAGCGTGGTCCGCGACT  663

Qy   727 GTGTGCCTTCCCCGGCGACTGCAGCTGCAGCTGGACTGGGAGCGGCTGAGCTCCGGC  786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664 GTGTGCCTTCCCCGGCGACTGCAGCTGCAGCTGGACTGGGAGCGGCTGAGCTCCGGC  723

Qy   787 TACGGCAAGCATGAGGCCTTGCTCCTTTCTCCTTTCATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC  846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   724 TACGGCAAGCATGAGGCCTTGCTCCTTTCTCCTTTCATTCGGAGCGGCGGCTGAAGGAGGCTCATGTC  783

Qy   847 AGACTGTACCATCCAGCCGCTGCACATCACAACACATTTACTTAACAGAACAGTCACCGAC  906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784 AGACTGTACCATCCAGCCGCTGCACATCACAACACATTTACTTAACAGAACAGTCACCGAC  843

Qy   907 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGAGCGCC  966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGAGCGCC  903

Qy   967 TGCCAGGGGCGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGGCCATGACTTTGGTG  1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   904 TGCCAGGGGCGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGGCCATGACTTTGGTG  963

Qy  1027 GGCATCATCAGCTGGGGGCCTGGGGCTGGGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1086
          ||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db   964 GGCATCATCAGCTGGGGGCCTGGGGCTGTGTGGACAGAAGGATGTCCCGGGTGTGTACACAAAG  1023

Qy  1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG  1128
          |||||||||||||||||||||||||||||||||||||||||||
Db  1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG  1065
```

RESULT 14
US-08-811-949-50
; Sequence 50, Application US/08811949
; Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

```
TELEPHONE: 703-413-3000
TELEFAX:  703-413-2220
INFORMATION FOR SEQ ID NO:  38:
SEQUENCE CHARACTERISTICS:
  LENGTH:  1974 base pairs
  TYPE:  nucleic acid
  STRANDEDNESS:  double
  TOPOLOGY:  circular
MOLECULE TYPE:  DNA (genomic)
FEATURE:
  NAME/KEY:  CDS
  LOCATION:  3..1583
US-08-811-949-38

Query Match          94.1%;   Score 1061.8;  DB 2;  Length 1974;
Best Local Similarity 99.3%;   Pred. No. 7e-289;
Matches 1066;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;
```

```
Qy     56  CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTACCGTG  115
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511  CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTTGGGAATGGGTCAGCCTACCGTG  570

Qy    116  GCACGCAGACGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA  175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571  GCACGCAGACGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA  630

Qy    176  TAGGCAAGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA  235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    631  TAGGCAAGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA  690

Qy    236  ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    691  ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  750

Qy    296  GGCTGACGTGGGAGTACTGGTGTCCCTCCTGTCTCAACCTGCGGCCTGAGACAGTACA  355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    751  GGCTGACGTGGGAGTACTGGTGTCCCTCCTGTCTCAACCTGCGGCCTGAGACAGTACA  810

Qy    356  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTCCACCCCTGGC  415
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    811  GCCAGCCTCAGTTTCGCATCATAGGAGGGCTTCTTCGCCGACATCGCCTCCACCCCTGGC  870

Qy    416  AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGAGAGGCGGTTCTTGTGCGGGGGCA  475
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    871  AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGGCGGTTCCTGTGCGGGGGCA  930

Qy    476  TACTCATCAGCTCCTGCTTGGATTCTCTGCCCGGCCCACTGCTTCTCCAGGAGAGGTTCCGC  535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    931  TACTCATCAGCTCCTGCTTGGATTCTCTGCCCGGCCCACTGCTTCTCCAGGAGAGGTTCCGC  990

Qy    536  CCCACCACCTGACGGTGATCTTCGGCAGAACATACCGGGTGGTCCTGGTCCTGGAGGAGAGC  595
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    991  CCCACCACCTGACGGTGATCTTCGGCAGAACATACCGGGTGGTCCTGGTCCTGGAGGAGAGC  1050

Qy    596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1051  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1110

Qy    656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGGCGGACTGCCCAGGAGCAGCG  715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1111  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGGCGGACTGCCCAGGAGCAGCG  1170

Qy    716  TGGTCCGGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171  TGGTCCGGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  1230

Qy    776  AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTTCGGAGCGGCGGCTGAAGG  835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231  AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTTCGGAGCGGCGGCTGAAGG  1290

Qy    836  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1291  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  1350
```

```
Qy    596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    779  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  838

Qy    656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGGCGGACTGCCCAGGAGCAGCG  715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGGCGGACTGCCCAGGAGCAGCG  898

Qy    716  TGGTCCGGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899  TGGTCCGGCACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG  958

Qy    776  AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959  AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  1018

Qy    836  AGGCTCATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACAACATTTACTTAACAGAA  895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019  AGGCTCATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACAACATTTACTTAACAGAA  1078

Qy    896  CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT  955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079  CAGTCACCGACAACACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT  1138

Qy    956  TGCACGACGCCTGCCAGGGCGATTCGGAGGGCCCCCTCGGTGTGTGTCTGAACGATGGCCCA  1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1139  TGCACGACGCCTGCCAGGGCGATTCGGAGGGCCCCCTCGGTGTGTGTCTGAACGATGGCCGCA  1198

Qy   1016  TGACTTTGGTGGGCATCATCAGCTGGGGCTCGGCTGTGAACAGAAGAGATGTCCCGGGTG  1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1199  TGACTTTGGTGGGCATCATCAGCTGGGGCTGGCCTGTGAACAGAAGAGATGTCCCGGGTG  1258

Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1259  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1311
```

```
RESULT 15
US-08-811-949-38
; Sequence 38, Application US/08811949
; Patent No. 5840533
GENERAL INFORMATION:
  APPLICANT:  NIWA, MINEO
  APPLICANT:  SAITO, YOSHIMASA
  APPLICANT:  SASAKI, HITOSHI
  APPLICANT:  HAYASHI, MASAKO
  APPLICANT:  NOTANI, JOUJI
  APPLICANT:  KOBAYASHI, MASAKAZU
  TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
  NUMBER OF SEQUENCES:  67
  CORRESPONDENCE ADDRESS:
    ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
    ADDRESSEE:  P.C.
    STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
    CITY:  ARLINGTON
    STATE:  VA
    COUNTRY:  USA
    ZIP:  22202
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/811,949
    FILING DATE:  05-MAR-1997
    CLASSIFICATION:  435
  ATTORNEY/AGENT INFORMATION:
    NAME:  OBLON, NORMAN F.
    REGISTRATION NUMBER:  24,618
    REFERENCE/DOCKET NUMBER:  18-966-0
  TELECOMMUNICATION INFORMATION:
```

```
Qy    896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1351 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1410

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1470

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGTG 1530

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1531 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1583
```

Search completed: August 19, 2003, 08:35:10
Job time : 89.2804 secs

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 318.307 Seconds
(without alignments)
7930.701 Million cell updates/sec

Title: US-09-987-455-5
Perfect score: 1128
Sequence: 1 atgaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 150479 seqs, 118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1128	100.0	1128	11	US-09-987-455-2	Sequence 2, Appli
2	1128	100.0	1128	11	US-09-987-455-5	Sequence 5, Appli
3	1063.4	94.3	1689	9	US-09-969-271-6	Sequence 6, Appli
4	1063.4	94.3	2509	14	US-10-193-656-7	Sequence 7, Appli
5	1063.4	94.3	2519	9	US-09-969-271-5	Sequence 5, Appli
6	1062	94.1	1065	11	US-09-987-455-4	Sequence 4, Appli
7	1062	94.1	1065	11	US-09-987-455-7	Sequence 7, Appli
8	1061.8	94.1	2641	10	US-09-974-298-144	Sequence 144, App
9	227.2	20.1	329	12	US-10-007-926A-433	Sequence 433, App
10	191.6	17.0	1212	10	US-09-880-503-15	Sequence 15, Appl
11	188	16.7	1415	9	US-10-198-846-12748	Sequence 12748, A
12	170.8	15.1	1475	10	US-09-735-705-122	Sequence 122, App
13	170.8	15.1	1475	10	US-09-850-716A-122	Sequence 122, App
14	170.8	15.1	1475	10	US-09-897-778-122	Sequence 122, App
15	170.8	15.1	1475	10	US-09-466-396A-122	Sequence 122, App
16	170.8	15.1	1475	12	US-10-117-982-122	Sequence 122, App
17	170.8	15.1	1475	12	US-10-101-510-159	Sequence 159, App
18	170.8	15.1	2336	12	US-09-971-392-46	Sequence 46, Appl
19	170.8	15.1	2336	12	US-10-101-510-458	Sequence 458, App
20	167.6	14.9	1296	13	US-10-076-421-1	Sequence 1, Appli
21	167.6	14.9	2294	9	US-09-735-705-123	Sequence 123, App
22	167.6	14.9	2294	10	US-09-850-716A-123	Sequence 123, App
23	167.6	14.9	2294	10	US-09-897-778-123	Sequence 123, App
24	167.6	14.9	2294	11	US-09-466-396A-123	Sequence 123, App
25	167.6	14.9	2294	12	US-10-117-982-123	Sequence 123, App
26	167.6	14.9	2344	12	US-10-101-510-383	Sequence 383, App
27	166	14.7	1236	10	US-09-880-503-12	Sequence 12, Appl
28	166	14.7	2304	12	US-10-301-822-160	Sequence 160, App
29	166	14.7	2304	14	US-10-171-311-183	Sequence 183, App
30	163	14.5	474	9	US-09-864-761-1682	Sequence 1682, Ap
31	162	14.4	2486	14	US-10-106-698-1989	Sequence 1989, Ap
32	159.8	14.2	831	10	US-09-880-503-14	Sequence 14, Appl
33	159.8	14.2	972	10	US-09-880-503-16	Sequence 16, Appl
34	139	12.3	461	11	US-09-918-995-10405	Sequence 10405, A
35	123	10.9	1959	12	US-10-101-510-743	Sequence 743, App
36	123	10.9	2035	11	US-10-101-510-448	Sequence 448, App
37	123	10.9	2048	11	US-09-858-909-1	Sequence 1, Appli
38	120.4	10.7	2036	10	US-09-954-456-552	Sequence 552, App
39	120.4	10.7	2036	10	US-09-880-107-1612	Sequence 1612, Ap
40	111.4	9.9	116	9	US-09-864-761-18438	Sequence 18438, A
41	105	9.3	457	11	US-09-918-995-11188	Sequence 11188, A
c 42	89.4	7.9	372	10	US-09-920-300A-237	Sequence 237, App
c 43	89.4	7.9	372	13	US-10-033-528-237	Sequence 237, App
44	88	7.8	442	11	US-09-918-995-12715	Sequence 12715, A
45	83.8	7.4	1230	9	US-09-879-792-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-987-455-2
; Sequence 2, Application US/0998745S
; Publication No. US2003004972SA1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match 100.0%; Score 1128; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACG 120

US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match   100.0%;   Score 1128;   DB 11;   Length 1128;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1128;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAAGACAGCTATCGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAAGACAGCTATCGCAGTTGCAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACG 120

Qy   121 CACAGCCTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGCCTCCCGTGGGAATTCCATGATCCTGATAGGC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACAGCCTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGCCTCCCGTGGGAATTCCATGATCCTGATAGGC 180

Qy   181 AAGGTTTACACAGCACAGAACCCCAGTGCCAAGCACTGGGCCTGGCAAACATAATTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAGGTTTACACAGCACAGAACCCCAGTGCCAAGCACTGGGCCTGGCAAACATAATTAC 240

Qy   241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCGCAGGCTG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCGCAGGCTG 300

Qy   301 ACGTGGGAGTACTGTGATGTGCCCTGCGCAGCTGGAGCAGAGCCAGAGTACAGCCAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACGTGGGAGTACTGTGATGTGCCCTGCGCAGCTGGAGCAGAGCCAGAGTACAGCCAG 360

Qy   361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTCGGCAGGCT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTCGGCAGGCT 420

Qy   421 GCCATCTTTGCCAAGCACAGAGGTCGCCCCGAGAGCGGCATACTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCATCTTTGCCAAGCACAGAGGTCGCCCCGAGAGCGGCATACTC 480

Qy   481 ATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGGAGAGTTCCGCCCCAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGGAGAGTTCCGCCCCAC 540

Qy   541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600

Db   541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600

Db   601 TTTGAAGTCGAAAAATACACCATTGTCCATAAGGAATTCGATGATGATGACACTTACGACAACAATGAC 660

Db   661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTC 720

Db   721 CGCACTGTGTGCCTTCCCCCGGGGACCTGCAGCTGGACGTGGAGCTC 780

Db   781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTTCTATTCGGAGCGGCTGAAGGAGGCT 840

Db   841 CATGTCAGACTGTACCCATCCAGCGCGCTGCACATCCAGCCACATCACAACAACAGTC 900

Db   901 ACCGACAACACATGCTGTGTGCTCCGAGCACTCCGAGCGCGGCGCCCCCAGGCAAACTTGCAC 960

Db   961 GACGCCTTGCCAGGCGGATTCGGGAGGGCTCGGGCTGGAGCTGTGTCTGAACGATGGCCGCATGACT 1020

Db 1021 TTGGTGGGCATCATCAGCTCGGGGCTGGGGCTGGGGCGTCCCGGGTGTGTAC 1080

Db 1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128

RESULT 2

```
Qy   236 ATTACTGCCGGAATCCTGATGGGATGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA   295
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   794 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA   853

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCTGCGGCCTGCGCGGGCCTGAGACAGTACA   355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   854 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCTGCGGCCTGCGCGGGCCTGAGACAGTACA   913

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCTGGC   415
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   914 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC   973

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCCTGTGCGGGGGCA   475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCCTGTGCGGGGGCA   1033

Qy   476 TACTCATCAGCTCCTGCTGATTCTCTGCTGCGCCCACTGCTGCTTCAGGAGAGGTTTCCGC   535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1034 TACTCATCAGCTCCTGCTGATTCTCTGCTGCGCCCACTGCTGCTTCAGGAGAGGTTTCCGC   1093

Qy   536 CCCACCACTGAGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGC   595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1094 CCCACCACTGAGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGC   1153

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA   655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA   1213

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATGGATTCGTCCGCTGTGCCCGGAGAGCAGCGG   715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1214 ATGACATTGCGCTGCTGCAGCTGAAATGGATTCGTCCGCTGTGCCCGGAGAGCAGCGG   1273

Qy   716 TGGTCCGGACCTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG   775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1274 TGGTCCGGACCTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG   1333

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTTTCATTCGGAGCGGCGGCTGAAGG   835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTTTCATTCGGAGCGGCGGCTGAAGG   1393

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCGTGCACATCACAACATTACTTAACAGAA   895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1394 AGGCTCATGTCAGACTGTACCATCCAGCCGCGTGCACATCACAACATTACTTAACAGAA   1453

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGCGCCCCCAGCAAACT   955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1454 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGCGCCCCCAGCAAACT   1513

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCGCCCTGTGTGTCTGAACGATGGCCGCA   1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1514 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCGCCCTGTGTGTCTGAACGATGGCCGCA   1573

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGCCATCATCAGCTGGGCCGA   1075
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1574 TGACTTTGGTGGGCATCATCAGCTGGGCCATCATCAGCTGGGCCGCA   1633

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG   1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1634 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG   1686
```

```
Qy   601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAGTTCGATGATGACACTTACGACAATGAC   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAGTTCGATGATGACACTTACGACAATGAC   660

Qy   661 ATTGCGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTC   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATTGCGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTC   720

Qy   721 CGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTC   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTC   780

Qy   781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCT   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCT   840

Qy   841 CATGTCAGACTGTACCCATCCAGCGCGCTGCACATCACAACATTACTTAACAGAACAGTC   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CATGTCAGACTGTACCCATCCAGCGCGCTGCACATCACAACATTACTTAACAGAACAGTC   900

Qy   901 ACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGGGGCGCCCCCAGCAAACTTGCAC   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 ACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGGGGCGCCCCCAGCAAACTTGCAC   960

Qy   961 GACGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACT   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GACGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGGCCGCATGACT   1020

Qy  1021 TTGGTGGGCATCATCAGCTGGGCCTGGACAGAAGGAGTGTCCCGGGTGTGTAC   1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TTGGTGGGCATCATCAGCTGGGCCTGGACAGAAGGAGTGTCCCGGGTGTGTAC   1080

Qy  1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG   1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG   1128
```

RESULT 3
US-09-969-271-6
; Sequence 6, Application US/09969271
; Patent No. US2002009179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PC510951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6

Query Match   94.3%;   Score 1063.4;   DB 9;   Length 1689;
Best Local Similarity   99.4%;   Pred. No. 0;
Matches 1067;   Conservative 0;   Mismatches 6;   Indels 0;   Gaps 0;

RESULT 4
US-10-193-656-7
; Sequence 7, Application US/10193656
; Publication No. US2003009673A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1.1577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656

```
Qy    56 CCCAGGGCCGGCCTCTGAGGGAAACAGTGACGCTACTTTGGGAATGGGTCAGCCTACCGTG   115
         ||| ||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 CCCCTGCCTGCTCTGAGGGAAACAGTGACGCTACTTTGGGAATGGGTCAGCCTACCGTG   673

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA   175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA   733

Qy   176 TAGGCAAGGTTACACAGCACAGAACCCCAGTGCCCAGCCACTGCCCTGGGCACTGGGCAAACATA   235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   734 TAGGCAAGGTTACACAGCACAGAACCCCAGTGCCCAGCCACTGCCCTGGGCACTGGGCAAACATA   793
```

; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X07393
; DATABASE ENTRY DATE: 1995-03-27
; RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match   94.3%;   Score 1063.4;   DB 14;   Length 2509;
Best Local Similarity 99.4%;   Pred. No. 0;
Matches 1067;   Conservative   0;   Mismatches   0;   Indels   6;   Gaps   0;

```
Qy    56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
         ||| ||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 726

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA 786

Qy   176 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 846

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCACGTGGCAGCAGCCGCA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCACGTGCCACGTGGCAGCAGCCGCA 906

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTGCGCGCCTGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTGCGCGCCTGAGACAGTACA 966

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 1026

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGCGTTCCTGTGCGGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1027 AGGCTGCCATCTTTGCCAAGCACAGGAGGCGTTCCTGTGCGGGGGCA 1086

Qy   476 TACTCATCAGCTCCTGCTTGGATTCTCTGCCCCACTGCTTCTCCAGGAGAGGTTTCCGC 535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1087 TACTCATCAGCTCCTGCTTGGATTCTCTGCCCCACTGCTTCTCCAGGAGAGGTTTCCGC 1146

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGGCGAGGAGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1147 CCCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGGCGAGGAGAGC 1206

Qy   596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGATGCACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1207 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGATGCACTTACGACA 1266

Qy   656 ATGACATTGCGCTGCTGAAATCGGATTCGTCCCGTGCGCCGGAGAGCAGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1267 ATGACATTGCGCTGCTGAAATCGGATTCGTCCCGTGCGCCGGAGAGCAGCG 1326

Qy   716 TGTTCCGCACTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1327 TGTTCCGCACTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1386

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTGTTCTCTTTATTCGAGCCGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

; CURRENT FILING DATE: 2002-07-10
(...)
US-10-193-656-7

```
Db  1387 AGCTCTCCGGCTACGGCAAGCATGAGGCCTGTTCTCTTTATTCGAGCCGCTGAAGG 1446

Qy   836 AGGCTCATGTCAGACTGTCAGACTGTACCCATCCAGCCGCGTCGCACATCACAACATCACATTTACTTACTTAACAGAA 895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1447 AGGCTCATGTCAGACTGTCAGACTGTACCCATCCAGCCGCGTCGCACATCACAACATCACATTTACTTACTTAACAGAA 1506

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1507 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACT 1566

Qy   956 TGCAGACGGCCTGCCAGGGCGATTCGGGAGGCGATTCGGGAGGCGCCCCTGTGTGTCTGAACGATGGCCGCA 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1567 TGCAGACGGCCTGCCAGGGCGATTCGGGAGGCGATTCGGGAGGCGCCCCTGTGTGTCTGAACGATGGCCGCA 1626

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCTGGTGGACAGAAGGATGTCCCGGGTG 1075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1627 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCTGGTGGACAGAAGGATGTCCCGGGTG 1686

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1687 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1739
```

RESULT 5
US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US2002009817A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match   94.3%;   Score 1063.4;   DB 9;   Length 2519;
Best Local Similarity 99.4%;   Pred. No. 0;
Matches 1067;   Conservative   0;   Mismatches   0;   Indels   6;   Gaps   0;

```
Qy    56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
         ||| ||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   690 CCCTGCCTGCTCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 749

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   750 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA 809

Qy   176 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCGCCTGGGCACTGGGCCTGGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCGCCTGGGCACTGGGCCTGGGCAAACATA 869

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCTGCTGAAGAACCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCTGCTGAAGAACCGCA 929

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTGCTCCACCTGCGCGGCCTGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTGCTCCACCTGCGCGGCCTGAGACAGTACA 989

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCCCTGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCCACCCCTGGC 1049

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGAGGCCGTTCCTGTGCGGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: coding
;   OTHER INFORMATION: sequence for K2S protein
US-09-987-455-4

Query Match         94.1%;  Score 1062; DB 11;  Length 1065;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1062;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     67 TCTGAGGGAAAACAGTGACTGACTTTGGGAATGGGTCAGCTTACCGTGCACGGCACAGC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTGAGGGAAAACAGTGACTGACTTTGGGAATGGGTCAGCCTTACCGTGCACGGCACAGC 60

Qy    127 CTCACCGAGCTGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCACCGAGCTGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120

Qy    187 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180

Qy    247 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGGCTGACGTGG 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGGCTGACGTGG 240

Qy    307 GAGTACTGTGATGTGCCCTGCTGCCTCCACCTGCGCAGGGCTGCCATC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGTACTGTGATGTGCCCTGCTGCCTCCACCTGCGCAGGGCTGCCATC 300

Qy    367 TTTCGCATCAAAGAGAGGGCTCTTCGTGCCGCAGATGCGCGGTTCGTGTGCGGGGGCAATCAGC 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCGCATCAAAGAGAGGGCTCTTCGTGCCGCAGATGCGCGGTTCGTGTGCGGGGGCAATCAGC 360

Qy    427 TTTGCCAAGCACAGGAGTCGCCGCGGAGAGCGGGTTTCCGCCCACACTCATCAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTGCCAAGCACAGGAGTCGCCGCGGAGAGCGGGTTTCCGCCCACACTCATCAGC 420

Qy    487 TCCTGCTGGATTCTTCTCTGCCGCAGGAGGAGTTCCGGAGGAGCGCCACCTG 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCTGCTGGATTCTTCTCTGCCGCAGGAGGAGTTCCGGAGGAGCGCCACCTG 480

Qy    547 ACGGTGATCTTGGGCAGACATACCGGGTGGTCCCTGGGCAGGCAGAAATTTGAA 606
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGTGATCTTGGGCAGACATACCGGGTGGTCCCTGGGCAGGCAGAAATTTGAA 540

Qy    607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600

Qy    667 CTGCTGCAGCTGAAATCGGATTCCCGCGTGTCCCGTGCCCAGGAGAGCAGCGGTCCGCACT 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGCTGCAGCTGAAATCGGATTCCCGCGTGTCCCGTGCCCAGGAGAGCAGCGGTCCGCACT 660

Qy    727 GTGTGGCCTTCCCCGGCCGGACCTGCAGCTGACTGACGGAGTGTGAGCTCTCCGGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTGTGGCCTTCCCCGGCCGGACCTGCAGCTGACTGACGGAGTGTGAGCTCTCCGGC 720

Qy    787 TACGGCAAGCAGCATGAGGCCTTGCTGTCTCCTTTCTATTCGGAGGAGGGCTCATGTC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TACGGCAAGCAGCATGAGGCCTTGCTGTCTCCTTTCTATTCGGAGGAGGGCTCATGTC 780

Qy    847 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACTAACAGTCACCGAC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTAACTAACAGTCACCGAC 840

Qy    907 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCC 900

Qy    967 TGCCAGGGCGATTCGGGAGGCCCCTGGTGTCGTCGAACGATGGCCCATGACTTTGGTG 1026
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TGCCAGGGCGATTCGGGAGGCCCCTGGTGTCGTCGAACGATGGCCCATGACTTTGGTG 960

Db   1050 AGGCTGCCATCTTTGCCAAGCACAGGAGGACGTCGGCCCGGAGAGCGGCGTTCCTGTGCCGGGGGCA 1109
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCGCCCACTGCTTCCCGGAGGAGGTTTCCGC 535

Db   1110 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCGCCGCCCACTGCTTCCCGGAGGAGGTTTTCCGC 1169
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGTGGCGAGGAGC 595

Db   1170 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGTGGCGAGGAGC 1229
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    596 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655

Db   1230 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCCCGCGTGTCCCGTGCGCCAGGAGAGCAGCG 715

Db   1290 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCGCCAGGAGAGCAGCG 1349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    716 TGGTCCGCACTGTGTGCCTTCCCCGGCCGGGACCTGCAGCTGACTGGACGGAGTGTG 775

Db   1350 TGGTCCGCACTGTGTGCCTTCCCCGGCCGGGACCTGCAGCTGACTGGACGGAGTGTG 1409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    776 AGCTCTCCGGCTACGGCAAGCAGCATGAGGCCTTGCTGTCTCCTTTCTATTCGGAGG 835

Db   1410 AGCTCTCCGGCTACGGCAAGCAGCATGAGGCCTTGCTGTCTCCTTTCTATTCGGAGG 1469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895

Db   1470 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT 955

Db   1530 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT 1589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    956 TGCACGAGCCTGCCAGGGCGATTCGGGAGGCCCCTGGTGTCGTGAACGATGGCCCGCA 1015

Db   1590 TGCACGAGCCTGCCAGGGCGATTCGGGAGGCCCCTGGTGTCGTGAACGATGGCCCGCA 1649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1016 TGACTTTGGTGGGGCATCATCAGCTGGGCTGGGGCTGGGGCTGTGTGAACGATGTCCCGGGTG 1075

Db   1650 TGACTTTGGTGGGGCATCATCAGCTGGGCTGGGGCTGGGGCTGTGTGAACGATGTCCCGGGTG 1709
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGCGACCG 1128

Db   1710 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGCGACCG 1762
```

RESULT 6
US-09-987-455-4
; Sequence 4, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
;   APPLICANT: Aranya Manosroi
;   APPLICANT: Jiradej Manosroi
;   APPLICANT: Chatchai Tayapiwatana
;   APPLICANT: Friedrich Goetz
;   APPLICANT: Rolf-Guenther Werner
;   TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;   TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;   FILE REFERENCE: 0652.2190001
;   CURRENT APPLICATION NUMBER: US/09/987,455
;   CURRENT FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: 60/268,574
;   PRIOR FILING DATE: 2001-02-15
;   PRIOR APPLICATION NUMBER: GB 0027779.8
;   PRIOR FILING DATE: 2000-11-14
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
;   LENGTH: 1065
;   TYPE: DNA
;   ORGANISM: Artificial Sequence

Qy 421 TCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db

Qy 547 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGGAGGAGGAGGCAGAAATTTGAA 606
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGGAGGCAGAAATTTGAA 540

Qy 607 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATGACATTGCG 666
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATGACATTGCG 600

Qy 667 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGTGGTCCGCACT 726
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGACGAGCGTGGTCCGCACT 660

Qy 727 GTGTGCCTTCCCCGGCGACCTGCAGCTGCGACCTGCGCCGGACTGCGACGAGTGTGAGCTCCGGC 786
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GTGTGCCTTCCCCGGCGACCTGCAGCTGCGACCTGCGCCGGACTGCGACGAGTGTGAGCTCCGGC 720

Qy 787 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 846
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780

Qy 847 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 906
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840

Qy 907 AACATGCTGTGTCGTGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCGCC 966
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 AACATGCTGTGTCGTGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCGCC 900

Qy 967 TGGCAGGGCGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 1026
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 TGGCAGGGCGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTG 960

Qy 1027 GGCATCATCAGCTGGGGCTGGGCTGTGGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1086
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 GGCATCATCAGCTGGGGCTGGGCTGTGGTGACAGAAGGATGTCCCGGGTGTGTACACAAAG 1020

Qy 1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
        ||||||||||||||||||||||||||||||||||||||||||
Db 1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1062

RESULT 7
US-09-987-455-7
; Sequence 7, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-7

Query Match   94.1%; Score 1062; DB 11; Length 1065;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1062; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   67 TCTGAGGGAAACAGTGACTGCTACTTTGGAATGGGTCAGCCTACCGTGGCACGCACAGC 126
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60

Qy 127 CTCACCGGAGTCGGGTGCGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCACCGGAGTCGGGTGCGCCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120

Qy 187 TACAGCACAGAACCCCAGTGCCCAGGCACTGCCTGGGCACTGGGGCAAACATAATTACTGCCGG 246
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TACAGCACAGAACCCCAGTGCCCAGGCACTGCCTGGGCACTGGGGCAAACATAATTACTGCCGG 180

Qy 247 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 306
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGG 240

Qy 307 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGGGCGTCGAGACAGTAGCAGCCAGCCTCCAG 366
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGGGCGTCGAGACAGTAGCAGCCAGCCTCCAG 300

Qy 367 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGCGCCTCCACCCCTGGCAGGCTGCCATC 426
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 TTTCGCATCAAAGGAGGGGCTCTTCGCCGACATGCGCCTCCACCCCTGGCAGGCTGCCATC 360

Qy 427 TTTGCCAAGCACAGGAGGTCGCCCGAGAGCCGGTTCCTGTGCTGCGGGGGCATACTCATCAGC 486
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TTTGCCAAGCACAGGAGGTCGCCCGAGAGCCGGTTCCTGTGCTGCGGGGGCATACTCATCAGC 420

Qy 487 TCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTG 546
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db

RESULT 8
US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002015626 3A1 1001470CB1
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

Query Match   94.1%; Score 1061.8; DB 10; Length 2641;
Best Local Similarity   99.3%; Pred. No. 0;
Matches 1066; Conservative   0; Mismatches   7; Indels   0; Gaps   0;

RESULT 9
US-10-007-926A-433
; Sequence 433, Application US/10007926A
; Publication No. US20030143539A1
GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 433
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (82)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (159)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (264)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (278)
; OTHER INFORMATION: a, t, c or g
US-10-007-926A-433

Query Match   20.1%;   Score 227.2; DB 12;   Length 329;
Best Local Similarity 97.4%;   Pred. No. 5.9e-62;
Matches 229; Conservative   0; Mismatches   6;   Indels   0;   Gaps   0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
              ||| ||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   810 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 869

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGA 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCGTGGAATTCCATGATCCTGA 929

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 989

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 1049

Qy   296 GGCTGAGGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 GGCTGAGGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 1109

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTTGGC 415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 1169

Qy   416 AGGCTGCCATCTTTGCCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
              ||||||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db   1170 AGGCTGCCATCTTTGCCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1229

Qy   476 TACTCATCAGCTCCTGCTGGAATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 TACTCATCAGCTCCTGCTGGAATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTCCGC 1289

Qy   536 CCCACCACTGACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGGCGAGGAGAGC 595
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 CCCACCACTGACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGGCGAGGAGAGC 1349

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1409

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1469

Qy   716 TGGTCCGGACTGTGTGCCTTCCCCGGGGACCTGAGCCTGCCGGAGGTGGACGGAGTGTG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 TGGTCCGGACTGTGTGCCTTCCCCGGGGACCTGAGCCTGCCGGAGGTGGACGGAGTGTG 1529

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGCCTTGTCTCCTTTCATTCGGAGCCGCTGAAGG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 AGCTCTCCGGCTACGGCAAGCATGAGCCTTGTCTCCTTTCATTCGGAGCCGCTGAAGG 1589

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1649

Qy   896 CAGTGCACGACAACATGCTGCCAGGCGGCAGACACTCGGAGCGCCCCCGGTCTGTCTGAAACT 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 CAGTGCACGACAACATGCTGCCAGGCGGCAGACACTCGGAGCGCCCCCGGTCTGTCTGAAACT 1709

Qy   956 TGCACGACGCCTGCCAGGGGGATTCGGGAGGGCCCCCGTGTGTGTGGGTCTGGACAGGATGGCCGCA 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1710 TGCACGACGCCTGCCAGGGGGATTCGGGAGGGCCCCCGTGTGTGTGGGTCTGGACAGGATGGCCGCA 1769

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGGCCTGGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1770 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGGCCTGGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1829

Qy   1076 TGTACACAAAGGTTACCAACTACCTACTAGACTCGATTCGTGACAACATGCGACCG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1830 TGTGTACACAAAGGTTACCAACTACCTACTAGACTCGATTCGTGACAACATGCGACCG 1882

RESULT 10
US-09-880-503-15
; Sequence 15, Application US/09880503
; Patent No. US20020119964A1

GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT:-.HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-880-503-15

Query Match   17.0%; Score 191.6; DB 10; Length 1212;
Best Local Similarity   52.3%; Pred. No. 2.1e-50;
Matches 549; Conservative   0; Mismatches 459; Indels 42; Gaps 4;

```
Qy    84 CTGCTACTTTGGGAATGGTCAGCCTACCGTGCACGCACAGCCTCACCGAGTCGGGTGC 143
Db   147 CTGCTATGAGGGGGATGGTCACTTTTACCGAGGAAAGGCCAGCACTGACACCATGGGCGG 206

Qy   144 CTCCTGCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACAGCACAGAACCC 203
Db   207 GCCCTGCCTGCCCCTGGAACTCTGCCCACTGTCCTTCAGCAAACGTACCATGCCCACAGATC 266

Qy   204 CAGTGCCCAGGCACTGGGCGCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263
Db   267 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCAGGAACCCAGACAACCGGAG 326

Qy   264 CAAGCCCTGGTGCCACGTGCTGAAGAACCCAGGCTGACGTGGGAGTACTGTGATGTGCC 323
Db   327 GCGACCCTGGTGTATGTGCAGTGGGCCTAAGCCGTTGTCCAAGAGTGCATGGTGCA 386

Qy   324 CTCCTCCACCTGCGGGCTGAGAGCAGTACAGCCA--------GCCTCAGTT 368
Db   387 TGACTGCGCAGATGGAAAATTAAAATTTCAGTGTGGCCAAAGACTCTGAGGCCCCGCTT 446

Qy   369 TCGGCATCAAAGGAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCGAGGCTGCCATCTT 428
Db   447 TAAGATTATTGGGGAGGAGAATTCACCACCATCGAGAACCAGCCCTGGTTTGGAGGCAGCCTCATCAGCCC 506

Qy   429 TGCCAAGCACAGGAGAGGTGCGCCCGGAGAGGCGGTTCCTGTGCGGGGGCATACTCATCAGCTC 488
Db   507 CAGGAGGCACCGGGGGGCTCTGTCA---CCTACGTGTGTGGAGGCAGCCTCATCAGCGCC 563

Qy   489 CTGCTCGGATTCTCTCTCGCCGGCCCACTGCTTCCAGGAGGAGGTTTCCGGCCACCACCTGAC 548
Db   564 TTGCTGGGGTGATCAGCGCCACACACTGCTTCATTGATTACCCAAAGAAGGAGGACTACAT 623

Qy   549 GGTGATCTTGGGCAGAACATACCGGGTGGTGTCCCTGGCGAGGAGGGCAGAAATTTGAAGT 608
Db   624 CGTCTACCTGGGTCGCTCGCCTCAAGGCTTAACTCCAACACGCAAGGGAGATGAAGTTTGAGGT 683

Qy   609 CGAAAAATACATTGTCCATAAGGAATTCGATGATGACA------CTTACGACAATGACAT 662
Db   684 GGAAAACCTTCATCCTCACAACAAGGACTACAGCCGCTGACACGCTTCCTCACCACACGACACAT 743

Qy   663 TGGCGCTGCTGCAGCTGAAATCGGATTCGTCGCCCGAGGAGCAGCCTGGTCCG 722
Db   744 TGCCTTGCTGCGAGATCCCGTTCCAAGGAGGGCAGGGTGCGCAGCCATCCCGGACTATACA 803

Qy   723 CACTGTGTGCCTTCCCCCGGCGACCTGTCCGGCGAGCTGCCCAGCGTGTGAGCGTCTC 782
Db   804 GACCATCTGCCTGCCCCTCGAGTGATAAACGATCCCAGTTTGGCCACAACGTGTCGAGATCAC 863

Qy   783 CGGCTACGGCAAGCATGAGGCCTTGCTCTCCTTTCTATTCGGCAAGAGGCGCTGAAAATGACTGT 842
Db   864 TGGCTTTGGAAAAGAGAGAATTCTACCGACTATCTCTATCCGACTATCTCTATTCCGGCAAGCAGCAGCTGAAAATGACTGT
```

```
Qy   843 TGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCAC 902
Db   924 TGTGAAGCTGATTTCCCACCGGGAGTGTCAGCAGCCCACTACTACGGCTCTGAAGTCAC 983

Qy   903 CGACAACATGCTGTGTGCGGAGACACTCGGAGCGGGGGCGGGCCCCAGGCAAACTTGCACGA 962
Db   984 CACCAAAATGCTATGTGCTGCTGCACCCC---------------AATGGAAAACAGA 1025

Qy   963 CGCCTGCCAGGGGCGATTCGGAGGGCCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTT 1022
Db  1026 TTCCTGCCAGGGAGACTCAGGGGGACCCCTGCTCGTCGTCCTCGTTCCCTCCAAGGCCGCATGACTTT 1085

Qy  1023 GGTGGGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAGGGATGTCCCGGGTGTGTACAC 1082
Db  1086 GACTGGAATTGTGAGCTGGGGCCCGTGGGATGTGCCCCTGGAAGGAGGACAGCCCAGGCGCTCACAC 1145

Qy  1083 AAAGGTTACCAACTACCACCTAGACTGGATTCG 1112
Db  1146 GAGAGTCTCACACTTCTTACCCTGGATCCG 1175
```

RESULT 11
US-10-198-846-12748
; Sequence 12748, Application US/10199846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12748
; LENGTH: 1415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 1414, 1415
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12748

Query Match   16.7%; Score 188; DB 14; Length 1415;
Best Local Similarity   91.1%; Pred. No. 3.1e-49;
Matches 255; Conservative   0; Mismatches 20; Indels 5; Gaps 5;

```
Qy    56 CCAGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
Db   895 CCCTGCCTGCTCTGAGGGAAACAGTGACTG-TACTTTGGGAATGGTCAGCCTACCGTG 953

Qy   116 GCACGCAAGCCTCACCGAGTCGGGTGCCTCCT-GCCTCCCGTGCTGGGGAATTCCATGATCCTG 174
Db   954 GCACGCACAG-CTCACCGAGTCGGGTGCCTCCTCCCGTGCCTCCGGAATTCCATGATCCTG 1012

Qy   175 ATAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCCACTGCCCGGGCTGGGTCCG 234
Db  1013 ATAGGCAAGGTTTACACAGCACAGAACCCCAGTGGGGACTACTACA 1072

Qy   235 AATTACTGCCGGAATCCTGATGGGGATGTCCGGCAAGCC 293
Db  1073 AAATTACTGCCGAAT-CTGATGGGGATGCTGAACGAACCG 1131

Qy   294 CAGGCTGACGTGGGAGTACTGTGATGTGCCCCTCCC 333
```

```
Db   1132 CAGGCTGACGTGGGAGTACTGTGATGTGGCCTCTGTCACC 1171

RESULT 12
US-09-735-705-122
; Sequence 122, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: McNeill, Patricia D.
;   APPLICANT: Fanger, Neil
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C14
;   CURRENT APPLICATION NUMBER: US/09/735,705
;   CURRENT FILING DATE: 2000-12-12
;   NUMBER OF SEQ ID NOS: 419
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-735-705-122

Query Match         15.1%;  Score 170.8;  DB 9;  Length 1475;
Best Local Similarity  51.3%;  Pred. No. 9.6e-44;
Matches    551;  Conservative    0;  Mismatches 457;  Indels  66;  Gaps  4;

Qy     84 CTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGCCTCACCGAGTCGGGTGC 143
          ||||||   |||||||||||      ||    |||  |||||| || ||||||||||
Db    287 CTGCTATGAGGGAATGGTCACTTTACCGAGGAAAGGCCAGCACTGACACACCATGGGCCG 346

Qy    144 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATCGTCCTTCCTCAGCAAACGTACCTACGT 203
          ||||||||    |||||||||
Db    347 GCCCTGCCTGCCCTGGGACTCTGCCACTGTCCTCCACTGCCCAAACGTACCACACAGATC 406

Qy    204 CAGTGCCCAGGCACTGGGCGCCTGGGCAAACATAATTACTGCCGGAATTACTGGGGATGC 263
          |||       |||||||||||      ||||          |||||||  ||| |||
Db    407 TGATGCTCTTCAGCTGGGCGCCTGGGAAACATAATTACTGCAGGACCCAGAACAACCGGAG 466

Qy    264 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGGGAGTACTGTGATGTGCC 323
          |||||||||||||   |||||||||  ||||| |||||| |||||||||||||||| |
Db    467 GCGACCCGGTGCTATGTGCAGGTGGGCCTAAAGCCGCTTGTCCAAGAGTGCATGGTGCA 526

Qy    324 CTCCTGTCTCCACCTGCGCGCCTGAG 347
          ||||||||   |
Db    527 TGACTGCGCAGATGCAGAAAAAAAGCCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGGACATCGCCTC 586

Qy    348 ---ACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTC 404
             |||||||      ||  |||  ||   |||   ||  |    ||||||||||
Db    587 CCAAAAGACTCTGAGGGCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGGAAGAA 646

Qy    405 CCACCCCTTGGCAGGCTGCCATCTTTGCCAAGCACAGAGGTCGCCCCGGAGACGGGTTCCT 464
          ||||||||||||||  ||    |||   ||||||||| |||  ||| ||||||||||||
Db    647 CCAGCCCCGGTTTGCGGCCATCCATCACAGGAGGCACCGGGG---GGGCTCTGCGTCCACCTACGT 703

Qy    465 GTGCGGGGCATACTCATCAGCTCCTGCTGGATTCCATGATCCGATCCTGCTTCCAGGA 524
          |||||||||     ||||||||| | |||||||| |  |||||||  |  || |||||
Db    704 GTGTGGAGGCAGCCTCATCAGCCCTGCTTGCGGTGATCAGCCGCCACACTGCTTCATTGA 763

Qy    525 GAGGTTTCCGCCCCACCACCTGCTGATCTTGGGCAGAACATACCGGGTGGTCCCTGG 584
          ||  |||||||||||||||       ||  |||| |||||||    ||  |||||||
Db    764 TTACCCAAAGAGGAGGAGACTACTACATCGTCTCACCTCCTGGGTCGCTCTAAAGGCTTAAGGCTTCAAGCTTAACTCCAACAC 823
```

```
Qy    585 CGAGGAGGAGCAGAAAATTTGAACTGCGAAAAATACATTGTCCATAAGGAATTCGATGATGA 644
          ||||||  |||  |||||||||||       |||  ||||||| |||||| |||||| |||
Db    824 GCAAGGGGAGATGAAGTTTGAGGTGGAAAACCTACCTCATCCTCACACAAGGACTACACGCGCTGA 883

Qy    645 CA------CTTACGACAATGACATTGCCTGCTGCACTGTGTCCAGCTGAAATCGGATTCGTCCGCTG 698
          ||      || |   |    ||| |||||| |||  ||| |||  ||  |||  ||||||||||
Db    884 CACGCTTGCTCACCACAACGACATTGCCTGCTGAAGATCCGTTCCAAGGAGGGCAGGTG 943

Qy    699 TGCCCAGGAGGACAGCCTGGTGGTCGCACTGTGTCGTCCTTCCCCGGCGGACCTGCAGCTGCC 758
          |||||||||||||||      |||||||     |||||     || |||||||||||||||
Db    944 TGGGCAGCCATCCGGACTATACAGACCATCTGCCTGCTGCCGATGTATAAACGATCCCCA 1003

Qy    759 GGACTGACGGAGCTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTA 818
          |||  |  ||| |     |     ||||||    |||  |||  |||      |||||||
Db   1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy    819 TTCGGAGCGGGCTGAAGGAGGCTCATGTGAAGCTGTACCCATCCAGCCGCTGCACATCACA 878
          |||||||| |   |||||     ||||||| || ||| |||||| ||||||||| |||||
Db   1064 TCCGGAGCAGCTGAAGATGACTGTTGTGAAGCTGATTTCCCACCGGGAGTGTCAGCAGCC 1123

Qy    879 ACATTACTTAACAGAACAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGG 938
          ||   ||||   |||||||    |||||||||| |||||||||||||    |||||||
Db   1124 CCACTACTACGGCTTCTGAAGTCACCACCAAAAATGCTGTGTGCTGACCCACAGTG--- 1180

Qy    939 CGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGGACCAGGGGAGCCCCCTGGTGTG 998
          ||            |  ||||||| |||||||||||||||||||| ||||||||||||
Db   1181 -------------GAAAACAGATTCCTGCCAGGGACTCAGGGGACCCCTCGTCTG 1225

Qy    999 TCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACA 1058
          |||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db   1226 TTCCCTCCAAGGCCGCATGACTTTGACTGGAATTGTGACTGACTTTGTGAGCTGGGATGTGCCCT 1285

Qy   1059 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATTCG 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1286 GAAGGACAAGCCAGGCGTCTACACGAGTCTCACACTTCTTACCCTGGATCCG 1339

RESULT 13
US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US2002011513QA1
; GENERAL INFORMATION:
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: McNeill, Patricia D.
;   APPLICANT: Retter, Marc W.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C15
;   CURRENT APPLICATION NUMBER: US/09/850,716A
;   CURRENT FILING DATE: 2001-05-07
;   NUMBER OF SEQ ID NOS: 440
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-850-716A-122

Query Match         15.1%;  Score 170.8;  DB 10;  Length 1475;
Best Local Similarity  51.3%;  Pred. No. 9.6e-44;
Matches    551;  Conservative    0;  Mismatches 457;  Indels  66;  Gaps  4;

Qy     84 CTGCTACTTTGGGAATGGGTCAGCCTACCGTGCACGCACAGCCTCACCGAGTCGGGTGC 143
          ||||||   |||||||||||      ||    |||  |||||| || |||||||||||
Db    287 CTGCTATGAGGGGAATGGTCACTTTACCGAGGAAAGGCCAGCACTGACACCATGGGCCG 346

Qy    144 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATCGTCCTTCCTCAGCAAACGTACCTACGT 203
          ||||||||    |||||||||
Db    347 GCCCTGCCTGCCCTGGGACTCTGCCACTGTCCTCCACTGCCCAAACGTACCACCACAGATC 406

Qy    204 CAGTGCCCAGGCACTGGGCGCCTGGGCAAACATAATTACTGCCGGAATTACTGCCGGGATGC 263
          |||       |||||||||||      ||||          |||||||  ||| |||
Db    407 TGATGCTCTTCAGCTGGGCGCCTGGGAAACATAATTACTGCAGGACCCAGAACAACCGGAG 466

Qy    264 CAAGCCCCTGGTGCCACGTGCTGCCAGGCCGCAGCGTGACGTGGGAGTACTGTGATGTGCC 323
          |||||||||||||   |||||||||  ||||| |||||| |||||||||||||||| |
Db    467 GCGACCCCGGTGCTATGTGCAGGTGGGCCTAAAGCCGCTTGTCCAAGAGTGCATGGTGCA 526

Qy    324 CTCCTGTCTCCACCTGCGCGCCTGAG 347
          ||||||||   |
Db    527 TGACTGCGCAGATGCAGAAAAAAAGCCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGGACATCGCCTC 586

Qy    348 ---ACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTC 404
             |||||||      ||  |||  ||   |||   ||  |    ||||||||||
Db    587 CCAAAAGACTCTGAGGGCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGGAAGAA 646

Qy    405 CCACCCCTTGGCAGGCTGCCATCTTTGCCAAGCACAGAGGTCGCCCCGGAGAGTCGGGTGC 464
          ||||||||||||||  ||    |||   ||||||||| |||  ||| ||||||||||||
Db    647 CCAGCCCCGGTTTGCGGCCATCCATCACAGGAGGCACCGGGG---GGGCTCTGCGTCCACCTACGT 703

Qy    465 GTGCGGGGCATACTCATCAGCTCCTGCTGGATTCCATGATCCGATCCTGCTTCCAGGA 524
          |||||||||     ||||||||| | |||||||| |  |||||||  |  || |||||
Db    704 GTGTGGAGGCAGCCTCATCAGCCCTGCTTGCGGTGATCAGCCGCCACACTGCTTCATTGA 763

Qy    525 GAGGTTTCCGCCCCACCACCTGCTGATCTTGGGCAGAACATACCGGGTGGTCCCTGG 584
          ||  |||||||||||||||       ||  |||| |||||||    ||  |||||||
Db    764 TTACCCAAAGAGGAGGAGACTACTACATCGTCTCACCTCCTGGGTCGCTCTAAAGGCTTAAGGCTTCAAGCTTAACTCCAACAC 823
```

```
; APPLICANT:  Henderson, Robert A.
; APPLICANT:  Peckham, David W.
; APPLICANT:  Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 457
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-897-778-122

Query Match       15.1%;    Score 170.8;   DB 10;   Length 1475;
Best Local Similarity   51.3%;   Pred. No. 9.6e-44;
Matches  551;  Conservative   0;  Mismatches 457;  Indels  66;  Gaps   4;
```

```
Qy     84 CTGCTACTTTGGGAATGGTCAGCTCAGCCTGCACGCACGCCTCACCGAGTCGGGTGC 143
             ||||| | ||||||  |||| ||||||||||||||| ||  |||||||||
Db    287 CTTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAGGCCAGCACCATGGGCGCG 346

Qy    144 CTCCTGCCTCCCGTGGAATTCCATGATCCGTCCACTGTGATAGGCAAGGTTACACAGCACAGAACCC 203
             ||| |||||||  ||||  | | |  |||||||||  ||   |  |||| ||   ||||
Db    347 GCCCTGCCTCCCTGGAACTCTGCCACTCTCGCCACTGTGCTTCAGCAAAGTACCCACCAGATC 406

Qy    204 CAGTGCCCAGGACTGGGGCTGGGCAACATAATTACTGCCGAAATCCTGATGGGGATGC 263
             |||   ||| ||  |||||||||| |||||||||||  || | ||  |  |  ||||
Db    407 TGATGCTCTTCAGCTGGGCTGGGCTGGGAAACATAATTACTACTGCAGGAACCCAGACAACCGGAG 466

Qy    264 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC 323
             |||||||||||  |  |||| || ||||  || ||   ||||||| | |||| |||||
Db    467 GCGACCCTGGTGTATGTGCAGGTTGGGGCCTAAAGCCGCTTGTCTGTGTCAAGAGTGCGATGGTGCA 526

Qy    324 CTCCTGCTTCCACCTGCGGCCTGAG------------------------------ 347
             ||||  ||||||||||||||||| |
Db    527 TGACTGCGCAGATGGAAGGGTCAGTTGGAAAAAGCCCTCTCTCCTCCAGAGAGATTAAAATTTCAGTGTGG 586

Qy    348 ---ACAGTACAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGGCCGACATCGCCTC 404
                 || |||||| ||| || | ||| | ||||||||| ||  |||  |||  |||
Db    587 CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGAGAA 646

Qy    405 CCACCCCTGGCAGGCTGCCATCTTGCCAAGACAGAGGTGCGCCCGGAGAGCGGTTCCT 464
             |||||||  ||||  ||||||||  |   |||  |  |||||||  ||||||| | |||
Db    647 CCAGCCCTGGTTGTGCGGGCCATCATCAGTGGGAATTGTCCCCGGAGAGCGGTTCCT 703

Qy    465 CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATCGAGAA 524
             |  |  |  ||| |||   |||||||  |   | ||||  ||||| | |  | || ||
Db    704 GTGTGAGGCAGCGTAGTGTGCAGGTGGGGGGTGATCAGCCCGTTGCTGTGGTGCTTCATTGA 763

Qy    525 GAGGTTTCCGCCCCACCACCTGAGCGGTGATCTTGGGCAGACAGAACATACCCGGGTGTCCCTGG 584
             || |||||| |||| || | | |||    |   | ||||| || || |   |||||| |  |||
Db    764 TTACCCAAAGAAGGAGGAGCAGCAGAAATTTGAAGTCGAAAAATACATTGTCGAAAAATACATCACA 823

Qy    585 CGAGGAGGACAGAAATTTGAAGTCAGAAAAATACATTGTCCATAAGGCTTAACTCCAACAC 644
             |||| || |||| |||| || || |||| | |||||||| |  ||  ||| || ||||||
Db    824 GCAAGGGGAGATGAAGTTTGAGGTGGAGATCACTGGCTTTGGAAAAGAGAATTCACCGACTATCTCTA 883

Qy    645 CA------CTTACGACAATGACATTGCCGTGCTGCTCAGCTGAGATCGGATTCGTCCCGCTG 698
             ||       ||  |  | |  | | | ||  |   |  |||| | || | | ||||| | |
Db    884 CACGCTTGCTCACCACAACGACAACGACAATTGTTTGGAAAAGAGAATTCTACCGACTATCTCTA 943

Qy    699 TGCCCAGGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCC 758
             |||| | |||||||| || ||||| |||||||||||||| | ||||| |||||| |
Db    944 TGGGCAGCCATCCCCGGACTATACAGACCATCTGCCTGCCCTCGATGTATAAACGATCCCCA 1003

Qy    759 GGACTGGACGGAGTGTGAGCTCTTCCGGCTACCTGCCTTGTCTCCTTCTA 818
             |||||||||| |||||||| ||| |  ||  ||   |||||||||||| ||
Db   1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGTGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy    819 TTCGGAGCGGCTGAAGGAGGAGGAGCAGCTGGGCCTGGGGGGCTGCCATCAGCTGCACATCACA 878
             |||||| |||||||||| |||||||||||  |||  |||||||||| |||||||| |   |||
Db   1064 TCCGGAGCAGCTGAAGATGACTGTGTTGTGAAGCTGATTTCCCACCGGAGTGTCAGCAGCC 1123

Qy    879 ACATTACTTAACAGAACAGTCACCGACAACAGTCTGTGTGCTGGAGACACTCGGAGCGG 938
             |||  |||| |  | || || |||| |||| |||||||||||| |||||| || | ||
Db   1124 CCACTACTACGGCTCTGAAGTCACCACCACCAAATGCTGTGTGCTGTGACCCCACAGTG-- 1180

Qy    939 CGGGCCCCAGGCAAACTTGCACGACGCCTGCGATTCGGCAGGGGATTCGGGAGGCCCTGGTGTG 998
             | |||||||| | || |||| |  |||| |||   || ||  || |  ||||| ||||| | ||
Db   1181 --------GAAAAACAGATTCCTGCCCAGGCAGGGGACTCAGGGGACCCCCGTCGTG 1225

Qy    999 TCTGAACGATGGCCCATGGACTTTGGTGGCATCAGCTGGGCATCAGCCTGGGGTGTGACA 1058
             | |||| || |||  |||| |||| || || |||||||| |   ||||||| |||||| |
Db   1226 TTCCCTCCAAGGCCGCATGGACTTTGGAGTTGACTGGAATGTGTGAGCTGGTGGGGATGTGCCCT 1285

Qy   1059 GAAGGATGTCCCGGGTGTGTACAAAAGTTACCAACTACCTGAATTCG 1112
             |||| |||||| ||| ||||| |||||  | ||  || |  ||||
Db   1286 GAAGGACAAGCCCAGGCCAGTCTCTACACGAGAGCTCTCACACTTCTCACACTTCTTGCCCTGGATCCG 1339
```

```
RESULT 14
US-09-897-778-122
; Sequence 122, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
```

```
RESULT 15
US-09-466-396A-122
; Sequence 122, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-122

Query Match            15.1%;   Score 170.8; DB 11; Length 1475;
Best Local Similarity  51.3%;   Pred. No. 9.6e-44;
Matches  551;  Conservative    0;  Mismatches  457;  Indels   66;  Gaps   4;
```

```
Qy     84 CTGCTACTTTGGAATGGGTCAGCTACCGTGGCACGCACCAGCCTCACCGAGTCGGGTGC 143
          ||||||  |||||| |||      |||  ||| |||||| ||  |||  ||||||
Db    287 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAGGCCAGCACTGACACCATGGCCG 346

Qy    144 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACAGCACAGACCC 203
          ||||||  ||
Db    347 GCCCTGCCTGCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCATGCCCACAGATC 406

Qy    204 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263
          |||||||| |||||||
Db    407 TGATGCTCTTCAGCTGGGCCTGGGCCTGGGGAAACATATTACTGCAGGAAACATACCGGAG 466

Qy    264 CAAGCCCTGGTGCCACGTGCTGCTGAAGAACCGCAGGTGGGAGTACTGTGTGATGTGCC 323
          |||||| |   |||||
Db    467 GCGACCCTGGTGCTATGTGCAGTGCAGTGGGCCTAAAGCGCGTGTCCAAGAGTGCATGGTCA 526

Qy    324 CTCCTGCTCCACCTGCGGCCTGAG------ 347
          ||||||  ||||   |||
Db    527 TGACTGCCGCAGATGGAAAAAAGCCCTCCTCCTCCAGAAGAATTAAAATTTCAGTGTGG 586

Qy    348 --ACAGTACAGCCAGCTCAGTTTGCCATGAAGGGGCTCTTCGCCCGACATGCCCTC 404
            |||   |||
Db    587 CCAAAAGACTCTGAGGCCCCCGCTTTAAGATTATTGGGGGAGAATTCACCACCATGCGAGAA 646

Qy    405 CCACCCCTTGGCAGCTGCCATCTTTGCCCAAGCACAGGAGGTCGCCCGGAGAGGGGTTCCT 464
          ||||   ||||
Db    647 CCAGCCCTTGGTTTGCGGCCATTCACGAGGAGGCACCGGGG---GGGCTCTGTCGTCCTTACGT 703

Qy    465 GTGCGGGGGCATACTCATCAGCTCCTGCTGGATTCTCCTCGCCGCGCCCACTGCTTCCAGGA 524
          |||  ||||| ||||    ||||||||||||   |||
Db   1064 TCCGGAGCAGCTGAAGATGACTGTGTGAAGCTGATTCCCACCGGAGTGTCAGCAGCC 1123

Qy    879 ACATTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGG 938
          ||||
Db   1124 CCACTACTACGGCTCTGAGTCACCACCAAAATGCTGTGTGCTGTGCTGACCCACAGTG--- 1180

Qy    939 CGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGCGATTCGGAGGCGCCCCTGGTGTG 998
          -------GAAAACAGATTCCTGCCAGGGACTCAGGGGACCCCTCGTCTG 1225
Db   1181 

Qy    999 TCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGGGCTGGTGTGGACA 1058
          |||| |     |||||
Db   1226 TTCCCTCCAAGGCCGCATGACTTTGACTGGAATTGTGAGCTGGCTGGGGCCGTGGATGTGCCCT 1285

Qy   1059 GAAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATTCG 1112
          ||||
Db   1286 GAAGGACAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCCTGGATCCG 1339
```

```
Db    704 GTGTGGAGGCAGCCTCATCAGCCCTTGCTGGGTGATCAGCGCGCCACACACTGCTTCATTGA 763
          |
Qy    525 GAGGTTCCGCCCACCACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGG 584
          ||||||   |||||||||||
Db    764 TTACCCAAAGAAGGAGGAGGACTACATCGTCTACCTGGGTCGCAAGGCTTAACTCCAACAC 823

Qy    585 CGAGGAGGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGA 644
          ||||||||     ||||||||
Db    824 GCAAGGGGAGATGAAGTTTGAGGTGGAAAACCTCATCCTTACACAAGGACTACAGCGGCTGA 883

Qy    645 CA------CTTACGACAATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTG 698
          ||       ||| ||   |||
Db    884 CACGCTTGCTCACCACCACGACCATCTGCCTGCGCCCTGCGTGATGTATAACGATCCCCA 943

Qy    699 TGCCCAGGAGCAGCGTGTCCGACTGTGTCCGGTTCCCCGGGGGACCTGCAGCTGCC 758
          |
Db    944 TGCGGAGCCATCCCGGACTATACAGACCAATCTGCCTGCCCTGCGATGTATAACGATCCCCA 1003

Qy    759 GGACTGGACCGGAGTGTGAGCTGCTGGCTTGGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 818
          ||||
Db   1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGTGAAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy    819 TTCGGAGCGGCTGAAGGAGCGCATGTCATGTCAGACTGTGAGCTGCACCATCACA 878
          ||||||| |||  |||
Db   1064 TCCGGAGCAGCTGAAGATGACTGTGTGAAGCTGATTCCCACCGGAGTGTCAGCAGCC 1123
```

Search completed: August 19, 2003, 14:22:27
Job time : 320.307 secs

```
                Copyright   GenCore version 5.1.6
                Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    August 18, 2003, 22:44:33 ; Search time 3299.95 Seconds
                                        (without alignments)
                                        8307.845 Million cell updates/sec

Title:          US-09-987-455-5
Perfect score:  1128
Sequence:       1 atgaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       22781392 seqs, 1215223805b residues

Total number of hits satisfying chosen parameters:     45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estcv:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                10: gb_est2:*
                11: gb_htc:*
                12: gb_est3:*
                13: gb_est4:*
                14: gb_est5:*
                15: em_estfun:*
                16: em_estom:*
                17: em_gss_hum:*
                18: em_gss_inv:*
                19: em_gss_pln:*
                20: em_gss_vrt:*
                21: em_gss_fun:*
                22: em_gss_mam:*
                23: em_gss_mus:*
                24: em_gss_pro:*
                25: em_gss_rod:*
                26: em_gss_phg:*
                27: em_gss_vrl:*
                28: gb_gss1:*
                29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	895	79.3	930	13	BX329047	BX329047 BX329047
2	853	75.6	912	13	BU149958	BU149958 AGENCOURT
3	850.2	75.4	962	13	BX389609	BX389609 AGENCOURT
4	818.6	72.6	891	13	BU146167	BU146167 AGENCOURT
5	813.8	72.1	993	13	BQ278177	BQ278177 AGENCOURT
6	811.2	71.9	883	13	BQ690978	BQ690978 AGENCOURT
7	807.6	71.6	918	13	BX389608	BX389608 AGENCOURT
8	796.2	70.6	930	13	BU845014	BU845014 AGENCOURT
9	782.8	69.4	888	13	BU157720	BU157720 AGENCOURT
10	765.2	67.8	950	13	BQ927899	BQ927899 AGENCOURT
11	764.4	67.8	916	12	BI765247	BI765247 603050338
12	706.8	62.7	948	10	BE616613	BE616613 60127886d
13	700.4	62.1	958	12	BG763582	BG763582 602735987
14	694.4	61.6	974	13	BQ960216	BQ960216 AGENCOURT
15	693.4	61.5	709	10	BE732704	BE732704 60157136b
16	688.2	61.0	1060	12	BM556999	BM556999 AGENCOURT
17	677	60.0	1042	10	BE732413	BE732413 601569644
18	672	59.6	957	13	BQ687538	BQ687538 AGENCOURT
19	665.8	59.0	948	13	BQ687779	BQ687779 AGENCOURT
20	665.2	59.0	704	9	AU134301	AU134301 AU134301
21	655.4	58.1	818	9	AU124602	AU124602 AU124602
22	655.2	58.1	903	13	BQ689095	BQ689095 AGENCOURT
23	653.2	57.9	830	13	BQ896938	BQ896938 AGENCOURT
24	653.2	57.9	867	13	BU179903	BU179903 AGENCOURT
25	652.2	57.8	904	13	BQ689840	BQ689840 AGENCOURT
26	652.2	57.8	911	13	BQ684734	BQ684734 AGENCOURT
27	651	57.7	818	12	BM006474	BM006474 603615409
28	649.6	57.6	797	13	BQ079911	BQ079911 AGENCOURT
29	642.2	56.9	1037	13	BU157493	BU157493 AGENCOURT
30	631.2	56.0	813	13	BG120780	BG120780 602348540
31	629.8	55.8	663	9	AU134463	AU134463 AU134463
32	624	55.3	624	10	BE549489	BE549489 601278502
33	585.4	51.9	587	9	BE394944	BE394944 601308225
34	561	49.7	561	9	AL047284	AL047284 DKFZp586H
35	553.4	49.1	555	14	CB128855	CB128855 K-EST0178
36	535.4	47.5	537	14	CB155052	CB155052 K-EST0213
37	534	47.3	964	12	BG769384	BG769384 602742678
38	520.4	46.1	905	13	BU859271	BU859271 AGENCOURT
39	514.8	45.6	883	14	CD359959	CD359959 AGENCOURT
40	514.4	45.6	516	14	CB140923	CB140923 K-EST0194
41	493.4	43.7	916	13	BU157241	BU157241 AGENCOURT
42	487.8	43.2	527	9	AW630119	AW630119 hh78e06.y
43	482.8	42.8	553	9	AW394127	AW394127 MR2-TT001
44	474.8	42.1	917	13	BU190904	BU190904 AGENCOURT
45	472.8	41.9	937	13	BQ679501	BQ679501 AGENCOURT

ALIGNMENTS

```
RESULT 1
BX329047/c
LOCUS        BX329047                 930 bp    mRNA      linear    EST 01-MAY-2003
DEFINITION   BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
             clone CSODI039YK15 3-PRIME, mRNA sequence.
ACCESSION    BX329047
VERSION      BX329047.1  GI:30307809
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 930)
  AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE      Full-length cDNA libraries and normalization
  JOURNAL    Unpublished
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 4498.r For
             more information about this cluster, see
             http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAI021ZD08_C
             S01980_1&cluster=4498.r. Contact : Feng Liang Email :
             fliang@lifetech.com URL : http://fulllength.invitrogen.com/
             InvitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID :
```

```
FEATURES
     source
              CS0BAI02IZD08_CS01980_1.
              1..930
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DI039YK15"
              /tissue_type="PLACENTA COT 25-NORMALIZED"
              /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
              /note="1st strand cDNA was primed with a NotI-oligo(dT)
              primer. Five prime end enriched, double-strand cDNA was
              digested with Not I and cloned into the Not I and EcoR V
              sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT   184 a      268 c      282 g      194 t      2 others
ORIGIN

Query Match          79.3%;  Score 895;  DB 13;  Length 930;
Best Local Similarity  98.6%;  Pred. No. 2.9e-222;
Matches 912;  Conservative  0;  Mismatches 12;  Indels 1;  Gaps  1;

Qy       67 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCACGCACAGC  126
            ||||||||||||||||||||||| |||||||||| ||||||||||| |||||||||
Db      927 TTTGAGGGAAACAGTGACTGCTAC-TTGGGAATGGGTCAGCCTACGTGGGCACGCAGC  869

Qy      127 CTCACCGAGTCGGGTGCCTTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGTT  186
            |||||||||||||||||||||||||| ||||||| |||||||||||||||||||||||||
Db      868 CTCACCGAGTCGGGTGCTTNCCTGCCTCCCTGAATTCCATGATCCTGATAGGCAAGTT  809

Qy      187 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  246
            ||||||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db      808 TACACAGCACAGACNCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGG  749

Qy      247 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCAGCCTGACGTGG  306
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db      748 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCAGCCTGACGTGG  689

Qy      307 GAGTACTGTGATGTGCCCCTCCTGTCTCCACCTGCGGGCCTGAGACAGTAGCCAGCCTCAG  366
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      688 GAGTACTGTGATGTGCCCCTCCTGTCTCCACCTGCGGGCCTGAGACAGTAGCCAGCCTCAG  629

Qy      367 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC  426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      628 TTTCGCATCAAAGGAGGGCTCTTCGCCCACATCGCCTCCCACCCCTGGCAGGCTGCCATC  569

Qy      427 TTTGCCAAGCACAGGAGGTGCGCCGGAGAGCGTGCGCGGGGGCATACTCATCAGC  486
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      568 TTTGCCAAGCACAGGAGGTGCGCCGGAGAGCGTGCGCGGGGGCATACTCATCAGC  509

Qy      487 TCCTGCTGGATTCTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG  546
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      508 TCCTGCTGGATTCTCTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTG  449

Qy      547 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGAATTTGAA  606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      448 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCAGGAGAATTTGAA  389

Qy      607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  666
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      388 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG  329

Qy      667 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCAGGAGAGCAGCGTGGTCCGGACT  726
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      328 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCAGGAGAGCAGCGTGGTCCGGACT  269

Qy      727 GTGTGCCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACGGAGTGTGAGCTCTCCGGC  786
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      268 GTGTGCCCTTCCCCCGGCCGGACCTGCAGCTGCCGGACGGAGTGTGAGCTCTCCGGC  209

Qy      787 TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGACGGCGCTGAAGGAGGCTCATGTC  846
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      208 TACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGACGGCGCTGAAGGAGGCTCATGTC  149

Qy      847 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  906
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      148 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC  89

Qy      907 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCGGGCCCCAGGCAAACTTGCACGACGCC  966
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCGGGCCCCAGGCAAACTTGCACGACGCC  29

Qy      967 TGCCAGGGCGATTCGGGAGGCCCCC  991
            |||||||||||||||||||||||||
Db       28 TGCCAGGGGGATTCGGGAGGCCCCC  4

RESULT 2
BU149958
LOCUS       BU149958                 912 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8137213 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184119 5', mRNA sequence.
ACCESSION   BU149958
VERSION     BU149958.1  GI:22663490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13572 row: m  column: 16
            High quality sequence start: 18
            High quality sequence stop: 644.
FEATURES             Location/Qualifiers
     source          1..912
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6184119"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site 2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGGAGCGGCCGCCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT    189 a      272 c      271 g      179 t      1 others
ORIGIN

Query Match          75.6%;  Score 853;  DB 13;  Length 912;
Best Local Similarity  98.9%;  Pred. No. 2.5e-211;
Matches 901;  Conservative  0;  Mismatches  5;  Indels 5;  Gaps  4;

Qy      197 AGAACCCCAGTGCCCAGGCACTGGG--CCTGGGCAAACATAATTACTGCCGGAATCCTGA  254
            |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        1 AGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGA  60

Qy      255 TGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCAGCCTGACGTGGGAGTACTG  314
```

AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen. This sequence belongs to sequence cluster
           4498.r For more information about this cluster, see
           http://www.genoscope.cns.fr/
           cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_3&cluster=4498.r.
           Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fullength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_3.
FEATURES   Location/Qualifiers
     source   1..962
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI039YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT   203 a   290 c   277 g   191 t   1 others
ORIGIN

Query Match   75.4%; Score 850.2; DB 13; Length 962;
Best Local Similarity 96.6%; Pred. No. 1.4e-210;
Matches 879; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

Qy 208 GCCAGGCACTGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGCCAAG 267
        |||||||||||| | |||||||| |||||||||||||||||||||||||||||||
Db   1 GCCAGGCACTGGCNGGCCTGGGCAAACTAATTACTGCCGGAATCCTGATGGGGATGCCAAG 60

Qy 268 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGAGTACTGTGTGTGCCCTCC 327
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGAGTACTGTGTGTGCCCTCC 120

Qy 328 TGTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGAGGGGCTC 387
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TGTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGAGGGGCTC 180

Qy 388 TTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCGCCATCTTTGCCAAGCACAGGAGGTCG 447
        ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 181 TTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCG 240

Qy 448 CCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCTGGATTCTCTGCC 507
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CCCGGAGAGCGGTTCCTGTGCGGGGGGCATACTCTGGATTCTCTGCC 300

Qy 508 GCCCACTGCTTCCAGGAGGAGGTTCCGCCCCCACCACCTGACGTGATCTTGGCAGAACA 567
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GCCCACTGCTTCCAGGAGGAGGTTCCGCCCCCACCACCTGACGTGATCTTGGCAGAACA 360

Qy 568 TACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCAT 627
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCAT 420

Qy 628 AAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCAGCTGAAATCGGAT 687
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 AAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCAGCTGAAATCGGAT 480

Qy 688 TCGTCCCGTGCCCAGAGAGCAGCAGCCGTGGTCGTCCCCGGGAC 747
        ||||||||||||||||||||||||||||||||||||||||||||
Db 481 TCGTCCCGTGCCCAGAGAGCAGCAGCCGTGGTCGTCCCCGGGAC 540

Qy 748 CTGCAGCTGCCGACTGACGAGTGTGAGCTTCCGGCTACGGCAAGCATGAGGCCTTG 807
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 CTGCAGCTGCCGACTGACGGAGTGTGAGCTTCCGGCTACGGCAAGCATGAGGCCTTG 600

Db   61 TGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGAGTACTG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 315 TGATGTGCCCTCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCAT 374

Db 121 TGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCAT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 375 CAAAGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATCTTTGCCAA 434

Db 181 CAAAGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATCTTTGCCAA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 435 GCACAGGAGGTCGCGCCGGAGAGCGGTCCATACTCATCAGCTCCTGCTG 494

Db 241 GCACAGGAGGTCGCCCCGGAGAGCGGTTCCGTTCTGTGCGGGGCATACTCATCAGCTCCTGCTG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 495 GATTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCCACCACCTGACGGTGAT 554

Db 301 GATTCTCTGCCGCCCACTGCTTCCAGGAGGAGTTTCCGCCCCCACCACCTGACGGTGAT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 555 CTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTGAAGTCGAAAA 614

Db 361 CTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTGAAGTCGAAAA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 615 ATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCA 674

Db 421 ATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCA 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 675 GCTGAAATCGGATTCGTCCCGCTGTGCCCGCTGTGCCCAGAGAGCAGCAGCCGTGGTCGTCCCACTGTGTCCT 734

Db 481 GCTGAAATCGGATTCGTCCCGCTGTGCCCGCTGTGCCCAGAGAGCAGCAGCCGTGGTCGTCCCACTGTGTCCT 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 735 TCCCCCGCGGCGGCACCTGCAGCTCGACGTGGACTGTGAGCTCCGGCTCTCCGGCTACGGCAA 794

Db 541 TCCCCCGCGGCGGCACCTGCAGCTCGACGTGGACTGTGAGCTCCGGCTCTCCGGCTACGGCAA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 795 GCATGAGGCCTTGCTGTTGTCTTCTATTCGGACCGCTGAAGGAGCGGCTCATGTCGACTGTA 854

Db 601 GCATGAGGCCTTGCTGTTGTCTTCTATTCGGACCGCTGAAGGAGCGGCTCATGTCGACTGTA 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 855 CCCATCCAGCCGCTGCACATCACAACATTTACTTAACGAAGAGTCACCGACAACATGT 914

Db 661 CCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACATTACCACGACAACATGT 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 915 GTGTGCTCGGAGACACTCGGAGCGCGGACACTCGGAGCGCGGACGCGCCCAGGG 974

Db 721 GTGTGCTCGGAGACACTCGGAGCGCGGACACTCGGAGCGCGGACGCGCCCAGGG 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 975 -CGATTCGGGAGG-CCCCGTGGTGTCTGAACGATGGCCGACGACTTGGTGGGGCATC 1032

Db 781 NCGATTCGGGAGGCCCCCCGTGGTGTCTGAACGATGGCCGCATGGGGCATC 840
        |||||||||||| ||||||||||||||||||||||||||
Qy 1033 ATCAGCTGGGGGCTGTGACAGAAGGATGT-CCCGGGTGTGTACACAAAGGTTAC 1091

Db 841 ATCAGCTGGGGGCCTGTGACAGAAGGATGTGCCCGGGTGTGTACACCAAGGTTAC 900
        ||||||||||||||||||||||||||||| |||||||||||||||||||||||
Qy 1092 CAACTACCTAG 1102
        |||||||||||
Db 901 AAACTACCTAG 911

RESULT 3
BX389609

LOCUS   BX389609   962 bp   mRNA   linear   EST 08-MAY-2003
DEFINITION BX389609 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389609
VERSION   BX389609.1 GI:30463470
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 962)

```
Qy    808  TCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGC  867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGC  660

Qy    868  TGCACATCACAACATTACTTAACAGACAGTCACCGACAACATGCTGTCTGCTGGAGAC    927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  TGCACATCACAACATTACTTAACAGACAGTCACCGACAACATGCTGTCTGCTGGAGAC    720

Qy    928  ACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGGGCGGATTCGGGAGGC  987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGGGCGGATTCGGGAGGC  780

Qy    988  CCCCTGGTGTGTCTGAACGATGCCGCATGACTTTGGTGGGGCATCATCAGCTGGGGGCCTG  1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  CCCCTGGTGTGTCTGAACGATGCCGCATGACTTTGGTGGGGCATCATCAGCCCTGGGGCCTG  840

Qy   1048  GGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACTAGACTGG  1107
           |||||||    |   |||    |||   |||   |||| |  |   |||||||||||
Db    841  GGCTGTGGAACAGAAGATGTCCCGGTTGTTCCCACCAAG--TTACCACTTACTAAACTGG  898

Qy   1108  ATTCGTGACA  1117
           |||| |||
Db    899  ATTCGGGACA  908
```

RESULT 4
BU146167
LOCUS   BU146167   891 bp   mRNA   linear   EST 03-SEP-2002
DEFINITION AGENCOURT 7981262 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6185661 5', mRNA sequence.
ACCESSION   BU146167
VERSION   BU146167.1 GI:22659699
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 891)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13576 row: m column: 22 .
FEATURES   Location/Qualifiers
     source   1..891
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6185661"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site 2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life

---

BASE COUNT   178 a   266 c   271 g   176 t
ORIGIN   Technologies."

Query Match   72.6%;   Score 818.6;   DB 13;   Length 891;
Best Local Similarity   98.2%;   Pred. No. 2.3e-202;
Matches 849;   Conservative   0;   Mismatches 14;   Indels   2;   Gaps   2;

```
Qy    202  CCCAGTGCCCAGGCACTGGGCCTGGGCCAAACATAATTACTGCGGAATTCCTGATGGGGAT  261
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  CCCAGTGCCCAGGCACTGGGCCTGGGCCAAACATAATTACTGCGGAATTCCTGATGGGGAT   60

Qy    262  GCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGGGAGTACTGTGATGTG   321
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GCCAAGCCCTGGTGCCACGTGCTGAAGAACCGAGGCTGACGTGGGAGTACTGTGATGTG   120

Qy    322  CCCTCCTGCTCACCTGCGGCTGAGCAGTCACAGCCAGCCTCAGTTTCGCATCAAAGGA   381
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CCCTCCTGTCCACCTGCGGCTGAGACAGTCACAGCCAGCCTCAGTTTCGCATCAAAGGA   180

Qy    382  GGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGCGTGCCATCTTTGCCAAGCACACAGG  441
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  GGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGCGTGCCATCTTTGCCAAGCACACAGG  240

Qy    442  AGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCCATACTCATCAGCTCCTGCTGCTGGATTCTC  501
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  AGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGCCATACTCATCAGCTCCTGCTGCTGGATTCTC  300

Qy    502  TCTGCGCCCACTGCTTCCAGGAGAGGTTTCGCCGGTGATCTTGGGC  561
           |||||||||||||||||||||||||||||||||||||||||||||
Db    301  TCTGCGCCCACTGCTTCCAGGAGAGGTTTCGCCGGTGATCTTGGGC  360

Qy    562  AGAACATACCGGTGGTCCTGCCGGAGGAGGAGGACAGAAATTTGAAGTCGAAAAATACATT  621
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGAACATACCGGTGGTCCTGCCGGAGGAGGAGGACAGAAATTTGAAGTCGAAAAATACATT  420

Qy    622  GTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCTGCAGCTGAAA  681
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTGCTGCTGCAGCTGAAA  480

Qy    682  TCGGATTCGTCCCGCGTGTCCCGCGTGTCCCGGAGGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCG  741
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TCGGATTCGTCCCGCGTGTCCCGCGTGTCCCGGAGGAGCAGCGTGGTCCGCACTGTGTGCCTTCCCCG  540

Qy    742  GCGGACTTGCAGCTGCCGGACTGGACGGAGTGGAGCTCTCCGGCTACGGCCAAGCATGAG  801
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GCGGACTTGCAGCTGCCGGACTGGACGGAGTGGAGCTCTCCGGCTACGGCCAAGCATGAG  600

Qy    802  GCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCC  861
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GCCTTGTCTCCTTTCTATTCGGAGCGGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCC  660

Qy    862  AGCCGCTGCACATCACAACATTACTTAACAGACAGTCACCGACAACAGTCGTGTGTGCT  921
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AGCCGCTGCACATCACAACATTACTTAACAGACAGTCACCGACAACAGTCGTGTGTGCT  720

Qy    922  GGAGACACTCGGAGCGGCGGC-GGGCCCCAGGCAAACTTGCCCAGGCGCAAGCGCCTGCCAGGGGCGGATTC  980
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    721  GGAGACACTCGGAGCGGCGGGCCCCCAGGCAAACTTGCCCAGGCGCAAGCGCCTGCCAGGGGCGCGATTC  780

Qy    981  GGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCAT-GACTTTGGTGGGGCATCATCAGCT  1039
           |||| |||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db    781  GGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTGGGGCATCATCAGCT  840

Qy   1040  GGGGCCTGGGCTGTGGACAGAAGGA  1064
           |||| |||||||||||||||||
Db    841  GGGGCCCCCTGGGCTGGGGGGACAGA  865
```

RESULT 5
BQ278177
LOCUS   BQ278177   993 bp   mRNA   linear   EST 07-MAY-2002
DEFINITION AGENCOURT_7061709 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5805153

```
ACCESSION   BQ278177
VERSION     BQ278177.1  GI:20488385
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 993)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2043  row: o  column: 10
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1..993
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5805153"
                     /tissue_type="teratocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
                     /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT     235 a    295 c    269 g    194 t
ORIGIN
Query Match       72.1%;  Score 813.8;  DB 13;  Length 993;
Best Local Similarity   99.6%;  Pred. No. 4.3e-201;
Matches  826;  Conservative   0;  Mismatches   2;  Indels  1;  Gaps   1;

Qy   300 GACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCTGAGACAGTACAGGCCA 359
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db     1 GACGTGGGAGTACTGTGATGTGCCCTCCTGCT-CACCTGCGGGCTGAGACAGTACAGGCCA  59

Qy   360 GCCTCAGTTCGCATCAAAGGAGGGCTTCGCCGGACATGCGCTTCCACCCCTGGCAGGC 419
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    60 GCCTCAGTTCGCATCAAAGGAGGGCTTCGCCGGACATGCGCTTCCACCCCTGGCAGGC 119

Qy   420 TGCCATCTTGCCAAGCACAGAGGTCGCCCCGAGAGGGGTTCCTGTCGCGGGGCATACT 479
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 TGCCATCTTGCCAAGCACAGAGGTCGCCCCGAGAGGGGTTCCTGTCGCGGGGCATACT 179

Qy   480 CATCAGCTCCTGCTGGAATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCA 539
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 CATCAGCTCCTGCTGGAATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCA 239

Qy   540 CCAACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCCGGAGGAGCAGAA 599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 CCAACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCCGGAGGAGCAGAA 299

Qy   600 ATTTGAAGTCGAAAAATACATTGTCCATAAAGGAATTCGATGATGACACTTACGACAATGA 659
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 ATTTGAAGTCGAAAAATACATTGTCCATAAAGGAATTCGATGATGACACTTACGACACAATGA 359

Qy   660 CATTGCGCTGCTGCAGTCGAAATCGGATTCGTCCCGCTGAATGGATTCGTCCCCAGGAGCAGCGTGGT 719
             |||||||||
```

```
Db   360 CATTGCGCTGCTGCAGTCGAAATCGGATTCGTCCCGCGTGCCCCAGGAGCAGCGGTGT 419

Qy   720 CCGACTGTGTGCCTTCCCCCGGGGACCTGCAGTCGCTGCCGACTGGACGGAGGTGTGAGCT 779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 CCGACTGTGTGCCTTCCCCCGGGGACCTGCAGTCGCTGCCGACTGGACGGAGGTGTGAGCT 479

Qy   780 CTCCGGCTACGGCAAGCATGAGCCTTGTCTCCTTCCTTCATTCGGAGCGGCTGAAGGAGGC 839
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   480 CTCCGGCTACGGCAAGCATGAGCCTTGTCTCCTTCCTTCATTCGGAGCGGCTGAAGGAGGC 539

Qy   840 TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTAACAGAACAGT 899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   540 TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTAACAGAACAGT 599

Qy   900 CACCGACACACATGTCGTGTCTGGAGACACTCGGAGCGGCGGGGCCCAGCCAAACTTGCA 959
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 CACCGACACACATGTCGTGTCTGGAGACACTCGGAGCGGCGGGGCCCAGCCAAACTTGCA 659

Qy   960 CGACGCCTGCCAGGGGGATTCGGAGGCCCCGTGGTGTCTGAACGATGGCCGCCATGAC 1019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 CGACGCCTGCCAGGGGATTCGGAGGCCCCGTGGTGTCTGAACGATGGCCGCCATGAC 719

Qy  1020 TTTGGTGGGCATCATCAGCTGGGGGCCTGGGCGTGGACAAGAAGGATGTCCCGGGTGTGTA 1079
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 TTTGGTGGGCATCATCAGCTGGGGGCCTGGGCGTGGACAAGAAGGATGTCCCGGGTGTGTA 779

Qy  1080 CACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACACATGCGGACACCG 1128
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 CACCAAGGTTACCAACTACCTAGACTGGATTCGTGACACACATGCGGACACCG 828
```

```
RESULT 6
BQ690978
LOCUS       BQ690978                 883 bp    mRNA          linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
            5', mRNA sequence.
ACCESSION   BQ690978
VERSION     BQ690978.1  GI:21816294
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2367  row: o  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1..883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6209120"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site 2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
```

```
                          (Stratagene) and Superscript II RT (Life Technologies).
                          Note: this is a NIH_MGC Library."

BASE COUNT       184 a     265 c     259 g     175 t
ORIGIN

Query Match      71.9%;  Score 811.2;  DB 13;  Length 883;
Best Local Similarity 99.0%;  Pred. No. 1.9e-200;
Matches 816;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy   178 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAAT 237
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAAT 60

Qy   238 TACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 297
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy   298 CTGACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGGGCCTGAGACAGTACAGC 357
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCCTGAGACAGTACAGC 180

Qy   358 CAGCCTCAGTTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAG 417
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGCCTCAGTTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAG 240

Qy   418 GCTGCCATCTTTGCCAAGCACACAGGAGGTGCGCCGGAGGCGGTTCCTGTGCGGGGGCATA 477
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCTGCCATCTTTGCCAAGCACACAGGAGGTGCGCCGGAGGCGGTTCCTGTGCGGGGGCATA 300

Qy   478 CTCATCAGCTCCTGCTGATTCTCTGCGGACTCCCACTGCTTCCAGGAGGTTTCCGCCC 537
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTCATCAGCTCCTGCTGATTCTCTGCGGACTCCCACTGCTTCCAGGAGGTTTCCGCCC 360

Qy   538 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCAGGAGGAGCAG 597
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGGTGGTCCCTGGCAGGAGGAGCAG 420

Qy   598 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 657
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 480

Qy   658 GACATTGGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTG 717
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GACATTGGGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTG 540

Qy   718 GTCCGACACTGTGTGCCTTCCCCCGGCGGACCTGAGCTGCGGGACTGGACGTGGTGTGAG 777
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTCCGACACTGTGTGCCTTCCCCCGGCGGACCTGAGCTGCGGGACTGGACGTGGTGTGAG 600

Qy   778 CTCTCCGGCTACGGCAAGCATGAGGGCCTTGCCTTTCTATTCGGAGGCGGCTGAAGGAG 837
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTCTCCGGCTACGGCAAGCATGAGGGCCTTGCCTTTCTATTCGGAGGCGGCTGAAGGAG 660

Qy   838 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACA 897
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACA 720

Qy   898 GTCACCGACAACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGCGGCCCCCAGGCAAACTTG 957
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GTCACCGACAACATGCTGTGTCGTGGAGACACTCGGAGCGGCGGCGGCCCCCAGGCAAACTTG 780

Qy   958 CACGAGCGCCTGCCAGGGGCGATTCGGGAGGCCCCCTGTGTGTGT 1001
         |||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACGAGCGCCTGCCAGGGGCGAATTCGGAGGCCCCCTGTGTGTGT 824


RESULT 7
BX389608

LOCUS       BX389608      918 bp    mRNA    linear    EST 08-MAY-2003
DEFINITION  BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389608
VERSION     BX389608.1   GI:30463469
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 918)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4498.r For more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_2&cluster=4498.r.
            Contact: Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_2.
FEATURES             Location/Qualifiers
     source          1..918
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI039YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT       189 a     273 c     264 g     186 t
ORIGIN

Query Match      71.6%;  Score 807.6;  DB 13;  Length 918;
Best Local Similarity 96.4%;  Pred. No. 1.7e-199;
Matches 855;  Conservative  0;  Mismatches 29;  Indels  3;  Gaps  3;

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    17 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 76

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCCGGCCTGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    77 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCCGGCCTGAGACAGTACA 136

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 GCCAGCCTCAGTTTCGCATCAAAGGAGGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 196

Qy   416 AGGCTGCCATCTTTGCCAAGCACACAGGAGGTCGCCCGGAGGAGGTCGCGCCGGAGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 AGGCTGCCATCTTTGCCAAGCACACAGGAGGTCGCCCGGAGGAGGTCGCGCCGGAGGGCA 256

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTGCGGACAGGGCTCTTCGCGGGGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 TACTCATCAGCTCCTGCTGGATTCTCTGCGGACAGGGCTCTTCGCGGGGTTTCCGC 316

Qy   536 CCCAGCACCTGACGGTGATCTTGGGCATCAAAGGAGGGGCTCTTCGCGGGGTGGGGAGAGC 595
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 CCCAGCACCTGACGGTGATCTTGGGCATCAAAGGAGGGGCTCTTCGCGGGGTGGGGAGAGC 376

Qy   596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 436

Qy   656 ATGACATTGCGCTGCTGCAGCTCGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 ATGACATTGCGCTGCTGCAGCTCGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 496

Qy   716 TGGTCCGACACTGTGTGCCTTCCCCCGGCGGCGGACCTGCAGCTGCGGGACTGCCGGAGTGTG 775
```

ORIGIN

Query Match   70.6%;   Score 796.2; DB 13;   Length 930;
Best Local Similarity   99.4%;   Pred. No. 1.6e-196;
Matches 820; Conservative   0; Mismatches   3; Indels   2; Gaps   2;

```
Qy    300 GACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCGGCTGAGACAGTACAGCCA 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCGGCTGAGACAGTACAGCCA 60

Qy    360 GCCTCAGTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGCCAGGC 419
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCCAGGC 120

Qy    420 TGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGGCGGTTCTCGTGCGGGGCATACT 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGGCGGTTCTCGTGCGGGGCATACT 180

Qy    480 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 539
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CATCAGCTCCTGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 240

Qy    540 CCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGGAGCAGAA 599
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGGAGCAGAA 300

Qy    600 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 659
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 360

Qy    660 CATTGCCGCTGCTGCAGCTGAAATCGGATTCGTTCCGCTGTGCCCAGGAGAGCAGCGTCGT 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATTCGCGCTGCTGCAGCTGAAATCGGATTCGTTCCGCTGTGCCCAGGAGAGCAGCGTCGT 420

Qy    720 CCGCACTGTGTGCCTTCCCCCGGACCTGCAGCTGACTGGACTGGACGGAGTGTGAGCT 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCGCACTGTGTGCCTTCCCCCGGACCTGCAGCTGACTGGACTGGACGGAGTGTGAGCT 480

Qy    780 CTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTCATTCGGAGGCGGCTGAGGAGGC 839
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTTCATTCGGAGGCGGCTGAGGAGGC 540

Qy    840 TCATGTCAGACTGTACCCATCCAGCCGCTGACATCACAACATTTACTTAACAGAACAGT 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGT 600

Qy    900 CACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCCCAGGCAAACTTGCA 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCCCAGGCAAACTTGCA 660

Qy    960 CGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGAC 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGAC 720

Qy   1020 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAACGATGTCCC-GGGTGTGT 1078
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    721 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAACGGATGTGTCCCGGGGTGTGT 780

Qy   1079 ACA-CAAAGGTTACCAACTACCTTAGACTGGATTCGTGACAACATG 1122
          ||| |||||||||||||||||||||||||||||||||||||||
Db    781 ACACCCAGGGTACCAACTACCTTAGACTGGATTCGTGACAACATG 825
```

RESULT 9
BU157720
LOCUS   BU157720   888 bp   mRNA   linear   EST 04-SEP-2002
DEFINITION AGENCOURT_8043192 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084476
            5', mRNA sequence.
ACCESSION   BU157720
VERSION   BU157720.1 GI:22671252
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens

```
Db    497 TGGTCCGACACTGTGTGCCTTCCCCCGGACCTGCAGCTGACTGGACGGAGTGTG 556
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTCATTCGGAGGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 AGCTCTCCGGCTACGGCAAGCATGAGGGCTTGTCTCCTTTCATTCGGAGGCGGCTGAAGG 616

Qy    836 AGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    617 AGCTCATGTCAGACTGTACCCATCATCCAGCCGCTGCACATCACAACATTTACTTAACANGA 676

Qy    896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    677 CAGTCACCGANCACATGCTGTGTGCTGGAGACACTCGNAGCGGCGGCCCCCA-GCAAACT 735

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 TGCACGAGCGCCTGCCAGGGCGATTNCGGAGGCCCCCTGGTGTGTCTTGACGATTGCCCAA 795

Qy   1016 TGA-CTTTGGTGGGCATCATCAGCTGGGG--CCTGGGCTGTGTGACAGAAGGATGTCCCGGG 1073
          ||| |||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db    796 TGACCTTTGGTGGCCATCATCAGCTGGGGCCCTGGGCTGTGTGACAGAAGGATGTTCCCGG 855

Qy   1074 TGTGTACACAAAGTTACCAACTACCTTAGACTGGATTCGTGACAACA 1120
          ||| || ||||||||||||||||||||||||||||||
Db    856 TGTGTCACCCAAGTTTACAACTACCTTAGACTGGATTCGGGACAACA 902
```

RESULT 8
BU845014
LOCUS   BU845014   930 bp   mRNA   linear   EST 16-OCT-2002
DEFINITION AGENCOURT_10412263 NIH_MGC_109 Homo sapiens cDNA clone
            IMAGE:6578533 5', mRNA sequence.
ACCESSION   BU845014
VERSION   BU845014.1 GI:24029455
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2780 row: g column: 13
            High quality sequence stop: 639.
FEATURES   Location/Qualifiers
     source   1..930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6578533"
                     /tissue_type="teratocarcinoma, cell line"
                     /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (life Technologies). Note: this is a
                     NIH_MGC Library."
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
BASE COUNT   199 a   279 c   266 g   186 t
```

REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             1  (bases 1 to 888)
TITLE        NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
             Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM2316 row: e   column: 21
             High quality sequence stop: 630.
FEATURES     Location/Qualifiers
  source     1..888
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6084476"
             /tissue_type="ductal carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC 110"
             /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
             Site_2: EcoRI; cDNA made by oligo-dT priming.
             Directionally cloned into EcoRI/XhoI sites using the
             following 5' adaptor: GGCACGAG(G). Library constructed by
             Ling Hong in the laboratory of Gerald M. Rubin (University
             of California, Berkeley) using ZAP-cDNA synthesis kit
             (Stratagene) and Superscript II RT (Life Technologies)."
             Note: this is a NIH_MGC Library."

BASE COUNT   181 a   265 c   270 g   172 t
ORIGIN

Query Match           69.4%;  Score 782.8;  DB 13;  Length 888;
Best Local Similarity 97.2%;  Pred. No. 4.9e-193;
Matches  861;  Conservative  0;  Mismatches  17;  Indels  8;  Gaps   6;

Qy    178 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGCCACTGGGCCTGGCAAACATAAT 237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGCCACTGGGCCTGGCAAACATAAT  60

Qy    238 TACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy    298 CTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGC 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGGGCCTGAGACAGTACAGC 180

Qy    358 CAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCCACCCCTGGCAG 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCCACCCCTGGCAG 240

Qy    418 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCATA 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCATA 300

Qy    478 CTCATCAGCTCCTGATTCTCTGCCGCCCACTGCTTCCAGGAGGAGGTTCCGCCC 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTCATCAGCTCCTGATTCTCTGCCGCCCACTGCTTCCAGGAGGAGGTTCCGCCC 360

Qy    538 CACCACCTGACCGTGACTTGGCAGAACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 597
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACCTGACCGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGCAG 420

Qy    598 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 480

Qy    658 GACATTGGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGAGCAGCGTG 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GACATTGGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGAGCAGCGTG 540

Qy    718 GTCCGCACTGTGTCCTTCCCCCGGCGACCTGCAGTGCCGGACCTGCCGGACTGCGGAGTGTGAG 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCCGCACTGTGTCCTTCCCCCGGCGACCTGCAGTGCCGGACTGCCGGACTGCGGAGTGTGAG 600

Qy    778 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTCCTTTCTATTCGGAGCGGCTGAAGGAG 837
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAG 660

Qy    838 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATTACTTACAGCAACA 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATTACTTACAGCAACA 720

Qy    898 GTCACCGACAACATGCTGT-GTGCTGAGACACTCGG--AGCGGCGGGCCCCAGGCAAAC 954
          |||||||||||||||||| |||||||||||||||||  |||||||||||||||||||||
Db    721 GTCACCGACAACATGCTGTGGTGTGGGAGACACTCGGGAGACACTCGGGGCCCCAGGCAAAC 780

Qy    955 TTGCACGACGCCT--GCCAGGGGGATTCGGAGGCCCCTGGT-GTGTCTGAACGATGG-C 1011
          |||||||||||||  |||||||||||||||||||||||||| ||||||||||||||| |
Db    781 TTGCACGACGCCTGCGCCGGGGGGGGATTCGGAGGCCCCTGGGTCTGAACGATGGCC 840

Qy   1012 CGGATGACTTT--GGTGGGCATCATCAGCTGGGGGCCTGGGGTCGTGG 1055
          ||||||||||   |||||||||||||||||||| || ||  |||
Db    841 CGGATGACTTTTGGGGGGCATCATCAACTGGGGGCCTGGGGTCGTGG 886

RESULT 10
BQ927899
LOCUS       BQ927899               950 bp    mRNA       linear    EST 20-AUG-2002
DEFINITION  AGENCOURT_8775988 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6373489
            5', mRNA sequence.
ACCESSION   BQ927899
VERSION     BQ927899.1  GI:22342930
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 950)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2550 row: p  column: 02
            High quality sequence start: 35
            High quality sequence stop: 684.
FEATURES    Location/Qualifiers
  source    1..950
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6373489"
            /tissue_type="large cell carcinoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_18"
            /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G).Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a

us-09-987-455-5.rst

Wed Aug 20 09:29:42 2003

---

SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 916)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11476 row: k  column: 09
            High quality sequence stop: 837.
FEATURES             Location/Qualifiers
     source          1..916
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5190656"
                     /lab_host="DH10B"
                     /clone_lib="NIH MGC 116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female. Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning). Average insert size 1.4 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber (Invitrogen). Research Genetics tracking code
                     023. Note: this is a NIH MGC Library."
BASE COUNT    204 a    272 c    256 g    184 t
ORIGIN

Query Match              Score 764.4;  DB 12;  Length 916;
Best Local Similarity  67.8%;  Pred. No. 3.1e-188;
Matches  776;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  1;

Qy  352 TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCC 411
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCC 60

Qy  412 TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGCACAGGAGTCGCCGGCGG 471
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGCACAGGAGTCGCCGGCGG 120

Qy  472 GGCATACTCATCAGCTCCTGCTGGATTCTCTGCGCTGCTTCTCTGCGCTGCTTCCAGGAGTTT 531
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GGCATACTCATCAGCTCCTGCTGGATTCTCTGCGCTGCTTCTCTGCGCTGCTTCCAGGAGTTT 180

Qy  532 CCGCCCCACCACCTGACCGGTGATCTTCGGGTCGGTGGTCCTGGCGAGAACATACCGGGAG 591
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCGCCCCACCACCTGACCGGTGATCTTCGGGTCGGTGGTCCTGGCGAGAACATACCGGGAG 240

Qy  592 GAGCAGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTAC 651
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GAGCAGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTAC 300

Qy  652 GACAATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCGCCCGTGTGCCCGGAGAGC 711
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GACAATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCGCCCGTGTGCCCGGAGAGC 360

Qy  712 AGCGTGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCGAGCTGCCGGACTGGACGGAG 771
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AGCGTGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCGAGCTGCCGGACTGGACGGAG 420

---

BASE COUNT    207 a    283 c    279 g    181 t
ORIGIN        NIH_MGC Library."

Query Match              Score 765.2;  DB 13;  Length 950;
Best Local Similarity  67.8%;  Pred. No. 2e-188;
Matches  806;  Conservative  0;  Mismatches  0;  Indels  3;  Gaps  3;

Qy  276 CCACGTGCTGAAGAACCCGAGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCAC 335
        |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db   43 CCCCGTGCTGAAGAACCCGACGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCAC 102

Qy  336 CTGCGGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGA 395
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  103 CTGCGGGCCTGAGACAGTACAGCCAGCCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGA 162

Qy  396 CATCGCCTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGCCGGAGA 455
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  163 CATCGCCTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGCCGGAGA 222

Qy  456 GCGGTTCCTGTGCGGGGGCGGGCATACTCATCAGCTCCTGCTGGATTCTCTGCGCCCACTG 515
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  223 GCGGTTCCTGTGCGGGGGGCGGGCATACTCATCAGCTCCTGCTGGATTCTCTGCGCCCACTG 282

Qy  516 CTTCCAGGAGAG-TTTCCGCCCCACCACCTGAGCGGTGATCTTGGGCAGAACATACCGGGG 574
        ||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  283 CTTCCAGGAGAGGTTTCCGCCCCACCACCCTGAGCGGTGATCTTGGGCAGAACATACCGGGG 342

Qy  575 TGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGCTGAAAAAATACATTGTCCATAAGAAT 634
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  343 TGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGCTGAAAAAATACATTGTCCATAAGAAT 402

Qy  635 TCGATGATGACACTTACGACACAATGACATTGCCGCTGCTGAGCTGAAATCGGATTCGTCCC 694
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  403 TCGATGATGACACTTACGACACAATGACATTGCCGCTGCTGAGCTGAAATCGGATTCGTCCC 462

Qy  695 GCTGTGCCCAGGAGAGCAGCCGTGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCGAGC 754
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  463 GCTGTGCCCAGGAGAGCAGCCGTGGTCCGCACTGTGTGCCTTCCCCGGCGGACCTGCGAGC 522

Qy  755 TGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTCACGGCGAAGCATGAGGCCTTGTCTCTT 814
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  523 TGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTCACGGCGAAGCATGAGGCCTTGTCCCCTT 582

Qy  815 TCTATTCGGAGCCGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACAT 874
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  583 TCTATTCGGAGCCGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACAT 642

Qy  875 CACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGA 934
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  643 CACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGA 702

Qy  935 GCGGCGGGGCCCCCAGGGCAAACTTGCACGACAGCGCCTGCCAGGGCGATTCGGGAGGCCCCCTGG 994
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  703 GCGGCGGGGCCCCCAGGGCAAACTTGCACGACAGCGCCTGCCAGGGCGATTCGGGAGGCCCCCTGG 762

Qy  995 TGTGTCTGAACGATGGCCGCATGACTTTGGTGGCATCATCAG-CTGGGGCGTGT 1053
        ||||||||||||||||||||||||||||||||||||||||||  |||||||||||
Db  763 TGTGTCTGAACGATGGCCGCATGACTTTGGTGGCATCATCAGCCTGGGGCGTGT 822

Qy 1054 GGACAG-AAGGATGTCCCGGGTGTGTACACAAAGTTACCAA 1094
        ||||||  ||| |  | |||||||||||||||||||||||
Db  823 GGACAGAGAAGAGGCCCCGGGTGGGGGACCCAGGGTACCCA 864

---

RESULT 11
BI765247
LOCUS       BI765247                 916 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603050338F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190656 5',
            mRNA sequence.
ACCESSION   BI765247
VERSION     BI765247.1  GI:15756825
KEYWORDS    EST.

```
Query Match              62.7%;  Score 706.8;  DB 10;  Length 948;
Best Local Similarity    93.3%;  Pred. No. 3.2e-173;
Matches 793;  Conservative  0;  Mismatches 52;  Indels 5;  Gaps 5;

Qy    72  GGAAAACAGTGACTGCTACT-TTGGGAATGGGTCAGCTCAGTGGGCACGCACCGTGGCACGCAGCCTCA  130
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     8  GAGAGACAGTGACTGCTATTATTGGGAATGGGTCAGCTCAGTGGGCACGCACCGTGGCACGCAGCCTCA   67

Qy   131  CCGAGTCGGGTGCCTCCTGCCTCCC-GTGGAATTCCATGATCCTGATAGGCAAGGTTTAC          189
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    68  CCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTAC           127

Qy   190  ACAGCACAGAACCCAGTGCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGGAAT              249
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128  ACAGCACAGAACCCAGTGCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGGAAT              187

Qy   250  CCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCGAAGAACCGCAGGCTGACGTGGGAG           309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188  CCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCGAAGAACCGCAGGCTGACGTGGGAG           247

Qy   310  TACTGTGATGTGCCCCTGCTCCACCTGCGGCGGCCTCCACCTGCGCAGGCTGCAGCTTCAGTTT       369
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248  TACTGTGATGTGCCCCTGCTCCACCTGCGGCGGCCTCCACCTGCGCAGGCTGCAGCTTCAGTTT       307

Qy   370  CGATCAAAGGAGGGGCTTCGGCAACATGCCTCCACCCCTGGCAGGCTGCATCTTT             429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308  CGATCAAAGGAGGGGCTTCGGCAACATGCCTCCACCCCTGGCAGGCTGCATCTTT             367

Qy   430  GCCAAGCACAGAGGTGCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCC        489
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368  GCCAAGCACAGAGGTGCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCC        427

Qy   490  TGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTGACG         549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428  TGCTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTGACG         487

Qy   550  GTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGCAGAAATTTGAAGTC           609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488  GTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGCAGAAATTTGAAGTC           547

Qy   610  GAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTG        669
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548  GAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTG        607

Qy   670  CTGCAGTCGAAATCGGATTCGTCCCGGTGCTGCCGGAGAGAGCAGCGGTGGTCCCCACTCGTG      729
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   608  CTGCAGTCGAAATCGGATTCGTCCCGGTGCTGCCGGAGAGAGCAGCGGTGGTCCGCACTGTG      667

Qy   730  TGCCTTCCCCCGGGACCTGCAGCTGCAGCTGGACGGAGTGTGAGCTGTGAGCTCTCGGCTAC      789
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668  TGCCTT-CCCCGCGGGACCTGCAGCTGCAGCTGGACGGAGTGTGAGCTGTGAGCTCTCGGCTAC     726

Qy   790  GGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCTGAAGGAGGCTCATGTCAGA          849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727  GGCAAGCATGAGG-CTTGTCTCCTTTCTATTCGGAGGCCGCTGAAGGGGCTCATGTCGGA        785

Qy   850  CTGTACCCATCCAGCCGCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGACAAC       909
          |||  ||||||||||||||||||||||||||||||||||||||| ||||  ||||||||||||
Db   786  CTG-GTCCATCCAGCCGCGCTGGGCTTCCCCACCTTCTTCAGGACAGCAGGCGGCCGGCCTGC     844

Qy   910  ATGCTGTGTG  919
          ||||| ||||
Db   845  TGGGGCTGGG  854
```

RESULT 12
BE616613

```
LOCUS       BE616613           948 bp   mRNA       linear   EST 24-AUG-2000
DEFINITION  6012788864F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3611007 5',
            mRNA sequence.
ACCESSION   BE616613
VERSION     BE616613.1  GI:9898212
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/,
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM267 row: p column: 16
            High quality sequence start: 2
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1..948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3611007"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH MGC 39"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California, Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life
                     Technologies)."
BASE COUNT      173 a    295 c    293 g    187 t
ORIGIN
```

RESULT 13
BG763582

```
LOCUS       BG763582           958 bp   mRNA       linear   EST 15-MAY-2001
DEFINITION  6027359877F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4861268 5',
            mRNA sequence.
ACCESSION   BG763582
VERSION     BG763582.1  GI:14074235
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 958)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1719  row: n  column: 21
            High quality sequence stop: 810.
FEATURES             Location/Qualifiers
     source          1..958
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4861268"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH MGC 49"
                     /note="Organ: sKin; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT   198 a   285 c   282 g   193 t
ORIGIN

Query Match       62.1%;  Score 700.4; DB 12; Length 958;
Best Local Similarity  93.8%;  Pred. No. 1.5e-17;
Matches 840; Conservative 0; Mismatches 41; Indels 15; Gaps 10;
```

```
Qy       56 CCCAGGCGGCCCTTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
            ||  || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 120

Qy      116 GCAGCGCACAGCCTCACCGAGTCGGGTGCCTCCTCGCCTCCCGTGGAATTCCATGATCCTGA 175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCGCCTCCCGTGGAATTCCATGATCCTGA 180

Qy      176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 240

Qy      236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCCAAGAACCGCA 295
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCCAAGAACCGCA 300

Qy      296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCGCCTGAGACAGTACA 355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGCTGACGTGGGGAGTACTGTGATGTGCCCTGCTCCACCTGCGCGCCTGAGACAGTACA 360

Qy      356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCAGCCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 420

Qy      416 AGGCTGCCATCTTTGCCAAGCACCAGGAGGTCGCCCGGAGAGGCGGTTCCTGTGCGGGGCA 475
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AGGCTGCCATCTTTGCCAAGCACCAGGAGGTCGCCCGGAGAGGCGGTTCCTGTGCGGGGCA 480

Qy      476 TACTCATCAGCTCCTGGAATTCTCTCGCCCGCCCCACTGCTTCCGCCGCTGCTTCTCCGCCAGGAGGGTTCCTG 535
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TACTCATCAGCTCCTGCTGGATTCTCTCGCCCGCCCCACTGCTTCCGCCGCTGCTTCTCCGCCAGGAGGGTTCCTG 540

Qy      536 CCCCACCACTGACGCGGTGATCTTGGGCAGCAGAACATACCGGGTGCCCTGGCGGGAGGAGGAGC 595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CCCACCACTGACGCGGTGATCTTGGGCAGCAGAACATACCGGGTGCCCTGGCGGGAGGAGGAGC 600

Qy      596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACA 655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACA 660

Qy      656 ATGACATTGCGCTGCTGACGTGAAATCGGATTCGT--CCGCTGTGCCAGGAGGAGCAGC 714
            ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      661 ATGACATTGCGCTGCTGACGTGAAATCGGATTCGTCCCCGCTGTGCCAGGAGGAGCAGC 720

Qy      715 -GTGGTCCGCACTGTGCCTT--CCCCGGCGGACCTGCAGCTGCCGACCTGGAC-GGA 770
             ||||||||||||||||||   |||||||||||||||||||||||||||||| |||
Db      721 GGTGGTCCGCACTGTGCCTTTCCCCCGGCGGACCTGCAGCTGCCGACCTGGACACGGGA 780

Qy      771 GTGTGAGCTCTCCG-GCTACGGCCAAGCAT--GAGGCCTTGTCCCTTTCTATTCGGAGCG 827
            |||||||||||||| |||||||||||||||  |||||||||||||||||||||||||||
Db      781 GTGTGAGCTCTCCGAGCTACGGCCAAGCATTGAGGGCCTTGTCCTTTCTATTCGGAGCG 840

Qy      828 GCTGAAGGAGGGCTCATGTCAGACTGTACCATCCAGCCGCGTGCACATCACAACATTTACT 887
            ||||||||||||||||||||||||||||||||||| || |||||||| ||||||||||||
Db      841 G-TGAAGGAGGGCTCATG--CGACTGGACCATTCACCGGTGGCCATTCACAACTTTACT 897

Qy      888 TAACAGAACAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGCGGGC 943
            ||||| |||| |||| ||||||||||| |||| |||||||||||| |||| |||||||
Db      898 T-ACAGAACAGTACC-- -GAAAATGGTGTGTGCGGGACCCCGACGGGGGCCCCAGGC 949
```

```
RESULT 14
BQ960216
LOCUS       BQ960216            974 bp    mRNA    linear    EST 21-AUG-2002
DEFINITION  AGENCOURT_8945604 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6459629
            5', mRNA sequence.
ACCESSION   BQ960216
VERSION     BQ960216.1  GI:22375694
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 974)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2645  row: a  column: 06
            High quality sequence stop: 603.
FEATURES             Location/Qualifiers
     source          1..974
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6459629"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH MGC 101"
                     /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
```

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT   215 a   293 c   279 g   187 t
ORIGIN

Query Match   61.6%;   Score 694.4;   DB 13;   Length 974;
Best Local Similarity 97.8%;   Pred. No. 5.5e-170;
Matches 757;   Conservative   0;   Mismatches 11;   Indels 6;   Gaps 5;

Qy   342 CCTGAGACAGTACAGCCAGCTCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGC 401
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 CCTGAGACAGTACAGCCAGCTCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGC 78

Qy   402 CTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGAGAGCGGTT 461
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 CTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGAGAGCGGTT 138

Qy   462 CCTGTGCGGGGGGCATACTCATCAGCTCCTGGATTCTCTCTGCCGCCCACTGCTTCCA 521
         ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   139 CCTGTGCGGGGGGCATACTCATCAGCTCCTGGATTCTCTCTGCCGCTC-CTGCTTCCA 197

Qy   522 GGAGAGGTTTCCGCCCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCC 581
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 GGAGAGGTTTCCGCCCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCC 257

Qy   582 TGGCGGAGGAGGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGA 641
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 TGGCGGAGGAGGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGA 317

Qy   642 TGACACTTACGACACAATGACATTGGCGCTGCTGCAGCTGAAATCGGATTCCCGCTGTGC 701
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 TGACACTTACGACACAATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGC 377

Qy   702 CCAGGAGAGCAGCGTGGTTCCGGACACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGA 761
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 CCAGGAGAGCAGCGTGGTTCCGGACACTGTGTGCCTTCCCCCGGCGGACCTGCAGCTGCCGGA 437

Qy   762 CTGGACGGAGTGTGAGCTCTCCGGCTACGACGCAAGCATGAGGGCCTTGTCTCCTTTCTATTC 821
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTGGACGGAGTGTGAGCTCTCCGGCTACGACGCAAGCATGAGGGCCTTGTCTCCTTTCTATTC 497

Qy   822 GGAGCGGCTGAAGGAGAGGCTCATGTGACCTGTACCAGACTGTCAGACTGTCACATCACAACA 881
         ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   498 GGAGCGGCTGAAGGAGAGGCTCATGTGACCTGTACCCATCCAGCCGCTGCACATCACAACA 557

Qy   882 TTTACTTAACAGAACAGTCACCGACACAATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGG 941
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 TTTACTTAACAGAACAGTCACCGACACAATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGG 617

Qy   942 GCCCCAGGCAAACTTGCACGACGACGCCTGCCAGGCGAGCGGATTCGGAGGAGG--CCCCTGGTGTGTC 1000
         |||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   618 GCCCCAGGCAAACTTGCACGACGACGCCTGCCAGGCGAGCGGATTCGGAGGAGGCCCCCCTGGTGTGTC 677

Qy 1001 TGAACGATGGCCCGCAT-GACTTTGGTGGGCATCATCAGCT-GGGGCCTGGGCGTGTGGACA 1058
         ||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||||||
Db   678 TGAACGATGGCCCGCATCATCAGCTTTGGTGGGCATCATCAGCTCGGGGCCTGGGCGTGTGGACA 737

Qy 1059 GAAGGATGTCCC--GGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATT 1110
         |||||||||||| |||||||||||||||||||||||||||||||||||||
Db   738 GAAGGATGCCCCCGGGGGGGGTACCCCCAGGGTTACCAACTACCTAAACTGGAAT 791

RESULT 15
BE732704
LOCUS   BE732704   709 bp   mRNA   linear   EST 15-SEP-2000
DEFINITION 601571366F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925600 5',
            mRNA sequence.
ACCESSION   BE732704
VERSION   BE732704.1 GI:10146696
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

FEATURES   Location/Qualifiers
     source   1..709
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3925600"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH MGC 21"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).   Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT   147 a   219 c   203 g   140 t
ORIGIN

Query Match   61.5%;   Score 693.4;   DB 10;   Length 709;
Best Local Similarity 99.7%;   Pred. No. 8.9e-170;
Matches 705;   Conservative   0;   Mismatches   1;   Indels   1;   Gaps   1;

Qy   263 CCAAGCCCTGGTGCCAGTGCTGAAGAACCGCAGGCTGCTGAAGAACCGCAGGCTGGGAGTACTGTGATGTGC 322
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCAAGCCCTGGTGCCAGTGCTGAAGAACCGCAGGCTGCTGAAGAACCGCAGGCTGTGATGTGC 60

Qy   323 CCTTCCGTGCTTCCACCTGCCGCGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAG 382
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCTTCCGTGCTTCCACCTGCCGCGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAG 120

Qy   383 GGCTCTTCGCCGACATGGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGA 442
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGCTCTTCGCCGACATGGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGA 180

Qy   443 GGTGCGCCCGAGAGCGGTTCCTGTGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCTGGATTCTCT 502
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGTGCGCCCGAGAGCGGTTCCTGTGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCCTGGATTCTCT 240

Qy   503 CTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCA 562
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCA 300

Qy   563 GAACATACCGGGTGGTCCCTGGCGGAGGAGGCAGAAATTTGAAGTCGAAAAATACATTG 622
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAACATACCGGGTGGTCCCTGGCGGAGGAGGCAGAAATTTGAAGTCGAAAAATACATTG 360

Qy   623 TCCATAAGGAATTCGATGATGACATTACGACACTTACGACATTGGCGCTGCTGCAGCTGCGAAT 682
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCATAAGGAATTCGATGATGACATTACGACACTTACGACATTGGCGCTGCTGCAGCTGCGAAT 420

Qy   683 CGGATTCGTCCCGCTGTGCCACTGTGGTCCGCACTGTGTCCTTCCCCCCGG 742
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGGATTCGTCCCGCTGTGCCACTGTGGTCCGCACTGTGTCCTTCCCCCCGG 480

```
Qy   743  CGGACCTGCAGCTGCCGGACTGACGGAGTGTGAGCTCTCCGCGCTACGGCAAGCATGAGG  802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGG  540

Qy   803  CCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCA  862
          ||||| |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CCTTGTCCCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCA  600

Qy   863  GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG  922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG  660

Qy   923  GAGACACTCGGACGCGGCGGCCCCAGGCAAACTTGCACGACGCCTGC  969
          |||||||||||||||||||||| |||||||||||||||||||||||
Db   661  GAGACACTCGGAGCGGCGGCGGG-CCCAGGCAAACTTGCACGACGCCTGC  706
```

Search completed: August 19, 2003, 08:29:18
Job time : 3300.95 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: August 18, 2003, 19:53:48 ; Search time 4736.31 Seconds
(without alignments)
9743.047 Million cell updates/sec

Title: US-09-987-455-2
Perfect score: 1128
Sequence: 1 atgaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1128	100.0	1128	6	AX431268	AX431268 Sequence
2	1128	100.0	1128	6	AX431271	AX431271 Sequence
3	1065	94.4	1314	6	A27435	A27435 DNA sequenc
4	1065	94.4	1314	6	E01937	E01937 Synthetic D
5	1065	94.4	1341	6	A30593	A30593 DNA for tis
6	1065	94.4	1419	6	A27725	A27725 DNA sequenc
7	1065	94.4	1419	6	AR059996	AR059996 Sequence
8	1065	94.4	1419	6	E01944	E01944 Synthetic D
9	1065	94.4	1689	6	I06609	I06609 Sequence 44
10	1065	94.4	1689	6	I06614	I06614 Sequence 46
11	1065	94.4	1689	6	I08789	I08789 Sequence 3
12	1065	94.4	1780	6	A06611	A06611 Synthetic n
13	1065	94.4	1780	6	A10226	A10226 Synthetic D
14	1065	94.4	1836	6	E01176	E01176 DNA encodin
15	1065	94.4	1836	6	I01581	I01581 Sequence 1
16	1065	94.4	1981	6	I01583	I01583 Sequence 3
17	1065	94.4	1989	6	I01256	I01256 Sequence 2
18	1065	94.4	2099	6	A07207	A07207 Artificial
19	1065	94.4	2100	6	A07197	A07197 H.sapiens m
20	1065	94.4	2100	6	A27429	A27429 cDNA sequen
21	1065	94.4	2100	6	E01934	E01934 DNA encodin
22	1065	94.4	2101	6	AR059986	AR059986 Sequence
23	1065	94.4	2151	6	E00654	E00654 cDNA encodi
24	1065	94.4	2162	6	AR017909	AR017909 Sequence
25	1065	94.4	2162	6	AR017931	AR017931 Sequence
26	1065	94.4	2162	6	I18796	I18796 Sequence 3
27	1065	94.4	2249	6	A07209	A07209 Artificial
28	1065	94.4	2544	6	I08530	I08530 Sequence 1
29	1065	94.4	2547	6	E01163	E01163 cDNA encodi
30	1065	94.4	2547	6	I07841	I07841 Sequence 1
31	1065	94.4	2547	6	I07989	I07989 Sequence 1
32	1065	94.4	2547	6	I09622	I09622 Sequence 1
33	1065	94.4	2549	6	E01055	E01055 DNA encodin
34	1065	94.4	2560	6	A01465	A01465 H.sapiens t
35	1065	94.4	7533	6	I08638	I08638 Sequence 2
36	1064	94.3	1689	6	I08637	I08637 Sequence 1
37	1063.4	94.3	1314	6	AR059989	AR059989 Sequence
38	1063.4	94.3	1506	6	E02247	E02247 cDNA sequen
39	1063.4	94.3	1689	6	A14927	A14927 tPA-DNA fro
40	1063.4	94.3	1689	6	A14930	A14930 tPA-DNA fro
41	1063.4	94.3	1689	6	A34204	A34204 Human t-PA
42	1063.4	94.3	1689	6	A34205	A34205 Human t-PA
43	1063.4	94.3	1689	6	AX427761	AX427761 Sequence
44	1063.4	94.3	1689	6	E02448	E02448 DNA encodin
45	1063.4	94.3	1689	6	E02663	E02663 DNA encodin

ALIGNMENTS

RESULT 1
AX431268
LOCUS AX431268 1128 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 2 from Patent WO0240650.
ACCESSION AX431268
VERSION AX431268.1 GI:21656150
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1
REFERENCE 1
AUTHORS Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
TITLE Methods for large scale production of recombinant dna-derived tpa
or k2s molecules
JOURNAL Patent: WO 0240650-A 2 23-MAY-2002;

Pred. No. is the number of results predicted by chance to have a

```
FEATURES
     source          1..1128
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT      238 a      332 c      332 g      226 t
ORIGIN

Query Match        100.0%;   Score 1128;  DB 6;  Length 1128;
Best Local Similarity  100.0%;   Pred. No. 5.5e-259;
Matches 1128;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  ATGAAAAAGACAGCTATCGCAGTGCAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCAG  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGAAAAAGACAGCTATCGCAGTGCAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCAG  60

Qy   61  GCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCTACCGTGGCACG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCTACCGTGGCACG 120

Qy  121  CACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCCAGGCACTGGGCCACGTGCTGGAAGCCTGGTGTGCCACGCCTGGAAGCCCTGGTG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  CACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCCAGGCACTGGGCCACGTGCTGGAAGCCTGGTGTGCCACGCCTGGAAGCCCTGGTG 180

Qy  181  AAGGTTTACACAGCACAGAACCCCAGTGCCAAGCCCTGGTGCGCCACGTGCTGCCACGTGCCCATGATCCTGAAACATAATTAC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  AAGGTTTACACAGCACAGAACCCCAGTGCCAAGCCCTGGTGCGCCACGTGCTGCCACGTGCCCATGATCCTGAAACATAATTAC 240

Qy  241  TGCCGGAATCCTGGGATCTTGGGCAGAGAACATACCGGGTGCCCTGGCAGGAGGAGCAGAAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TGCCGGAATCCTGGGATCTTGGGCAGAGAACATACCGGGTGCCCTGGCAGGAGGAGCAGAAA 300

Qy  301  ACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCGGCCTGAGACAGTACAGCCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  ACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCGGCCTGAGACAGTACAGCCAG 360

Qy  361  CCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCT 420

Qy  421  GCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCGGTTCGCTCCGCCCCAC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCGGTTCGCTCCGCCCCAC 480

Qy  481  ATCAGCTCCTGCTGGAATTCTCTCGCCGCCGCCCACTGCTTCCCGGAGGAGAGTTCGCTACCGTGGCTGGCCCAG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ATCAGCTCCTGCTGGAATTCTCTCGCCGCCGCCCACTGCTTCCCGGAGGAGAGTTCGCTACCGTGGCTGGCCCAG 540

Qy  541  CACCTGACGGTGATCTTTGGGCAGAGAACATACCGGGTGCCCTGGCAGGAGGAGCAGAAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  CACCTGACGGTGATCTTTGGGCAGAGAACATACCGGGTGCCCTGGCAGGAGGAGCAGAAA 600

Qy  601  TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAGACACTTACGACAATGAC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAGACACTTACGACAATGAC 660

Qy  661  ATTGCGCTGCTGCAGTGAAATCGGATTCGGCCGTGCCCGGACCAGGAGCAGCGTGGTC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  ATTGCGCTGCTGCAGTGAAATCGGATTCGGCCGTGCCCGGACCAGGAGCAGCGTGTC 720

Qy  721  CGCACTGTGTGCCTTCCCCGGCGGACCTGCCGGACTGGACGCTGCCGGAGTGTGAGCTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  CGCACTGTGTGCCTTCCCCGGCGGACCTGCCGGACTGGACGCTGCCGGAGTGTGAGCTC 780

Qy  781  TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCCGGCTGAAGGAGGCT 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCCGGCTGAAGGAGGCT 840

Qy  841  CATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  CATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTC 900
```

```
Qy  901  ACCGACAACATGCTGTGTGCTGATGAGACACTCGGAGGCGGCGGCGGCGGCCCCAGGCAAACTTGCAC  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  ACCGACAACATGCTGTGTGCTGATGAGACACTCGGAGGCGGCGGCGGCGGCCCCAGGCAAACTTGCAC  960

Qy  961  GACGCCTGCCAGGGCGATTCGGAGGCCCCCTCGTGTGTCTGAACGATGCCGCATGACT 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  GACGCCTGCCAGGGCGATTCGGAGGCCCCCTCGTGTGTCTGAACGATGCCGCATGACT 1020

Qy 1021  TTGGTGGGCATCATCAGCTGGGCGTGGGCGCTGGCGTGTGACAGAAGGATGTCCGGGTGTGTAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  TTGGTGGGCATCATCAGCTGGGCGTGGGCGCTGGCGTGTGACAGAAGGATGTCCGGGTGTGTAC 1080

Qy 1081  ACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACAACATGGCGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  ACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACAACATGGCGACCG 1128
```

```
RESULT 2
AX431271

LOCUS       AX431271                1128 bp    DNA     linear   PAT 28-JUN-2002
DEFINITION  Sequence 5 from Patent WO0240650.
ACCESSION   AX431271
VERSION     AX431271.1   GI:21656153
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE     Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
  JOURNAL   Patent: WO 0240650-A 5 23-MAY-2002;
            BOEHRINGER INGELHEIM INT (DE)
FEATURES
     source          1..1128
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT      238 a      332 c      332 g      226 t
ORIGIN

Query Match        100.0%;    Score 1128;  DB 6;  Length 1128;
Best Local Similarity  100.0%;    Pred. No. 5.5e-259;
Matches 1128;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  ATGAAAAAGACAGCTATCGCAGTGCAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCAG  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGAAAAAGACAGCTATCGCAGTGCAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCAG  60

Qy   61  GCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCTACCGTGGCACG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTGGCCTACCGTGGCACG 120

Qy  121  CACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCCAGGCACTGGGCCACGTGCTGGAAGCCTGGTGTGCCACGCCTGGAAGCCCTGGTG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  CACAGCCTCACCGAGTCGGGTGCCTCCTGCTCCTCCGTGCCCAGGCACTGGGCCACGTGCTGGAAGCCTGGTGTGCCACGCCTGGAAGCCCTGGTG 180
```

```
Db   361 CCTCAGTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTCCCACCCCTGGCGGCT 420
ORIGIN

Qy   421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGAGAGCCGGTTCTCGTGCGGGGCATACTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGAGAGCCGGTTCTCGTGCGGGGCATACTC 480

Query Match             94.4%;  Score 1065;  DB 6;  Length 1314;
Best Local Similarity   99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Mismatches 0;  Conservative 0;  Indels 5;  Gaps 0;

Qy    56 CCCAGGCGGCCTCTGAGCGAAACAGTGACTGCTACTTGGGATGGGTCAGCCTACCGTG 115
          | || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 CCCCTGCCTCTGAGGGAAACAGTGACTGCTACTTGGGATGGGTCAGCCTACCGTG 298

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCCGTGCCAGCACCCCAGTGCCAGGCACTGGGCCTCGGGCACTGCCAGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCAGGCACTGGGCCTCGGGCACTGCCAGGCAAACATA 358

Qy   116 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCCTCGGCGTGGTCCCTGGCAGAGGAGAGCAGAAA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCCTCGGCGTGGTCCCTGGCGAGGAGAGCAGAAA 600

Qy   176 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCCAGGCACTGGGCTGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCCAGGCACTGGGCTGGCAAACATA 418

Qy   236 ATTACTGCCGAATCCTGATGGGATGCCCAAGCCCTGGTGCCCACGTGCTGAGAACCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 ATTACTGCCGAATCCTGATGGGATGCCCAAGCCCTGGTGCCCACGTGCTGAGAACCGCA 478

Qy   296 GGCTGACGTGGGGAGTACTGTCGCCGGACCTGCCGGACTGCCGGAGTGTGAGCTC 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 GGCTGACGTGGGGAGTACTGTCGCCGGACCTGCCGGACTGCCGGAGTGTGAGCTC 538

Qy   356 GCCAGCCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGGACATCGCCTCCCACCCCTGGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 GCCAGCCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGGACATCGCCTCCCACCCCTGGC 598

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGGCA 658

Qy   476 TACTCATCAGCTCCTGCTCTGATTCTCTGCTCCAGGAGAGGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TACTCATCAGCTCCTGCTCTGATTCTCTGCTCCAGGAGAGGTTTCCGC 718

Qy   536 CCCCACCTGACGGTGATCTTGGGCAGAACATACGGGTGCCTGGGAGGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   719 CCCCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 778

Qy   596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   779 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 838

Qy   656 ATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGCCGGCGGACTGTGCCGCACTGCCGCAGCTGCCAGCGAGCAGCGAGCGCACCGAGTGTG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   839 ATGACATTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGCCGGCGGACTGTGCCGCACTGCCGCAGCTGCCAGCGAGCAGCGAGCGCACCGAGTGTG 898

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACTGCCAGCTGCAGCTGACGGAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   899 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACTGCCAGCTGCAGCTGACGGAGTGTG 958

Qy   776 AGCTCTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   959 AGCTCTCCGGCTACGGCCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGG 1018

Qy   836 AGGCTCATGTGACTGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1019 AGGCTCATGTGACTGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1078

Qy   896 CAGTCACCGACAACACGTGTGTCTGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1079 CAGTCACCGACAACACGTGTGTCTGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT 1138

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCCTGGTGTCTGAACGATGGCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1139 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCCTGGTGTCTGAACGATGGCCGCA 1198

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGCGTGGGCTGGGCTGGGCTGTGGACAGAAGATGTCCCGGGTG 1075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1199 TGACTTTGGTGGGCATCATCAGCTGGGCGTGGGCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1258

Qy  1021 TTGGTGGGCATCATCAGCTGGATTCGGATTCGTCGACAACATGGCCGGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GACGCCTGCCAGGGCGATTCGGGAGGCCGTGGCTGGTGTCTGGACAGAAGGATGTCCCGGGTGTGTAC 1080

Qy  1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGCCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGGCCGACCG 1128
```

RESULT 3

```
LOCUS       A27435                1314 bp    DNA     linear           PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pTQkPAdeltatrp.
ACCESSION   A27435
VERSION     A27435.1  GI:2172230
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1314)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 35 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1314
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      286 a    386 c    393 g    249 t
```

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1259 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1311

RESULT 4
E01937
LOCUS   E01937   1314 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION   E01937
VERSION   E01937.1 GI:2170185
KEYWORDS   JP 1989104167-A/4.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1314)
  AUTHORS   Niwa.M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
  JOURNAL   Patent: JP 1989104167-A 4 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT   OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1989104167-A/4
            PD   21-APR-1989
            PF   01-AUG-1988 JP 1988192320
            PR   03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
            13-NOV-1987 GB 87 8726683
            PI   NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
            PI   NOTANI JOJI,
            PI   KOBAYASHI MASAKAZU
            PC   C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
            C12R1:19),
            PC   (C12N9/64,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense:'No;
            CC   *source: clone=pTQkPAdeltatrp;
            FH   Key   Location/Qualifiers
            FH
            FT   CDS   1..1314
            FT   /product='t-PA'.
FEATURES   Location/Qualifiers
     source   1..1314
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT   286 a   386 c   393 g   249 t
ORIGIN

Query Match   94.4%; Score 1065; DB 6; Length 1314;
Best Local Similarity   99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   56 CCCAGGCGGCCGCTGCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 CCCCTGCCTGCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 298

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGCTCCGTTGGAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGCTCCGTGGAATTCCATGATCCTGA 358

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCGCTGGGCAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 TAGGCAAGGTTTACACAGCACAGCACAGAACCCCAGTGCCCAGGCACTGGGCGCTGGGCAAACATA 418

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGGAGAGAACCGCA 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGGAGAGAACCGCA 478

Qy   296 GGCTGACCTGGGAGTACTGTGATGTGCCCACCTGCTCCACCTGCGGCGGCCTGAGACAGTACA 355

Db   479 GGCTGACCTGGGAGTACTGTGATGTGCCCACCTGCTCCACCTGCGGCGGCCTGAGACAGTACA 538
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   356 GCCAGCCTCAGTTTGCATCAAAGAGGGCTCTTCGGCCACATCGCCTCCACCCCTGGC 415

Db   539 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGGCCACATCGCCTCCACCCCTGGC 598

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCA 658

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 718

Qy   536 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   719 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 778

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   779 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 838

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCCG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   839 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCCG 898

Qy   716 TGGTCCGACACTGTGTGCCTTCCCCGGCGACCTGACCTGACCAGCTGCCGACGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   899 TGGTCCGACACTGTGTGCCTTCCCCGGCGACCTGACCTGACCAGCTGCCGACGGAGTGTG 958

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   959 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGG 1018

Qy   836 AGGCTCATGTCGAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACAACAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 AGGCTCATGTCGAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTACAACAGAA 1078

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGCCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGCCAAACT 1138

Qy   956 TGCACGACGCCTGCCAGGGACGTTCGGGAGGCCCCCTGGTGTGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1139 TGCACGAGCGCCTGCCAGGGGATTCGGGAGGCCCCCCTGTGTGTGTCTGAACGATGGCCGCA 1198

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGAGCAGCTGGGGCTGGGGCTGGACGAAGAAGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1199 TGACTTTGGTGGGCATCATCAGCTGGAGCAGCTGGGGCTGGGGCTGGACAGAAGAAGATGTCCCGGGTG 1258

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1259 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1311

RESULT 5
A30593
LOCUS   A30593   1341 bp   DNA   linear   PAT 04-AUG-1995
DEFINITION DNA for tissue plasminogen activator (t-PA) from patent EP0400545.
ACCESSION   A30593
VERSION   A30593.1 GI:1249165
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1341)
  AUTHORS   Stern,A., Kohnert,U., Rudolph,R., Fischer,S. and Martin,U.
  TITLE   Derivative of tissue plasminogen activator
  JOURNAL   Patent: EP 0400545-A 2 05-DEC-1990;
            BOEHRINGER MANNHEIM GMBH
FEATURES   Location/Qualifiers
     source   1..1341

```
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:3630"
BASE COUNT      293 a    397 c    399 g    252 t
ORIGIN

Query Match     94.4%;   Score 1065;   DB 6;   Length 1341;
Best Local Similarity 99.5%;   Pred. No. 6.5e-244;
Matches 1068;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

Qy      56  CCCAGGCGGCCTCTGAGGGAAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG  115
            |||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     269  CCCCTGCCTGCTCTGAGGGAAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG  328

Qy     116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA  175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA  388

Qy     176  TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
            |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db     389  TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  448

Qy     236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     449  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  508

Qy     296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     509  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  568

Qy     356  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGC  415
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     569  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGC  628

Qy     416  AGGCTGCCATCTTTGCCAAGCACAGAGGTCGCCGGGAGAGCGGTTCGTCGTGTGGGGGCA  475
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     629  AGGCTGCCATCTTTGCCAAGCACAGAGGTCGCCGGGAGAGCGGTTCGTCGTGTGGGGGCA  688

Qy     476  TACTCATCAGCTCCGCTGGATTCTCTCGCCGCCACTGCTTCTCCGCCGACTGCTTTCCGC  535
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     689  TACTCATCAGCTCCGCTGGATTCTCTCGCCGCCACTGCTTCTCCGCCGACTGCTTTCCGC  748

Qy     536  CCCACCACCTGACGGTGATCTTTGGGCAGAAACATACCGGGTTGTCCCGGCAGAACATA  595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     749  CCCACCACCTGACGGTGATCTTTGGGCAGAAACATACCGGGTTGTCCCGGCAGAACATA  808

Qy     596  AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     809  AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  868

Qy     656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCAGCGTGTGCCCAGGAGCAGCG  715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     869  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCAGCGTGTGCCCAGGAGCAGCG  928

Qy     716  TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGAACTGGACCGGAGTGTG  775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     929  TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGAACTGGACCGGAGTGTG  988

Qy     776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG  835
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     989  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG  1048

Qy     836  AGGCTCATGTACCGATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  895
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1049  AGGCTCATGTACCGATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA  1108

Qy     896  CAGTCACCGACAACATGCTGTGTGCTGGAGACTCGGAGCGGCGGCCCCCAGGCAAACT  955
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1109  CAGTCACCGACAACATGCTGTGTGCTGGAGACTCGGAGCGGCGGCCCCCAGGCAAACT  1168

Qy     956  TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTTGTGTGTCTGACGATGGCCGCA  1015
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1169  TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTTGTGTGTCTGACGATGGCCGCA  1228
```

```
Qy    1016  TGACTTTGGTGGGCATCATCAGCTGTGGGCGCCTGGGCTGTGTGAACAGAAGGATGTCCCGGGTG  1075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1229  TGACTTTGGTGGGCATCATCAGCTGTGGGCGCCTGGGCTGTGTGAACAGAAGGATGTCCCGGGTG  1288

Qy    1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGCGACCG  1128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1289  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1341


RESULT 6
A27725
LOCUS       A27725                  1419 bp    DNA        linear   PAT 09-JUL-2002
DEFINITION  DNA sequence of coding region in pmTQk112.
ACCESSION   A27725
VERSION     A27725.1  GI:21727236
KEYWORDS    
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE     New tissue plasminogen activator
  JOURNAL   Patent: EP 0302456-A 49 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES             Location/Qualifiers
     source          1..1419
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT      309 a    413 c    426 g    271 t
ORIGIN

Query Match     94.4%;   Score 1065;   DB 6;   Length 1419;
Best Local Similarity 99.5%;   Pred. No. 6.5e-244;
Matches 1068;   Conservative   0;   Mismatches   5;   Indels   5;   Gaps   0;

Qy      56  CCCAGGCGGCCTCTGAGGGAAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG  115
            |||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     344  CCCCTGCCTGCTCTGAGGGAAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG  403

Qy     116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA  175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     404  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA  463

Qy     176  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     464  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  523

Qy     236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     524  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  583

Qy     296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     584  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  643

Qy     356  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGC  415
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     644  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGC  703

Qy     416  CCCAGGCGGCCTCTGAGGGAAAACAGTGGCTACTTTGGGAATGGTCAGCCTACCGTG  475
            |||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     629  CCCCTGCCTGCTCTGAGGGAAAACAGTGGCTACTTTGGGAATGGTCAGCCTACCGTG  688

Qy      56  CCCAGGCGGCCTCTGAGGGAAAACAGTGGCTACTTTGGGAATGGTCAGCCTACCGTG  115
            |||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     344  CCCCTGCCTGCTCTGAGGGAAAACAGTGGCTACTTTGGGAATGGTCAGCCTACCGTG  403

Qy     116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA  175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     404  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA  463

Qy     176  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     464  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  523

Qy     236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     524  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  583

Qy     296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     584  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGGCCTGAGACAGTACA  643

Qy     356  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGC  415
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     644  GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTGCCGACATCGCCTCCCACCCCTGGC  703

Qy     416  AGGCTGCCATCTTTGCCAAGCACAGAGGTCGCCGGGAGAGCGGTTCGTCGTGTGGGGGCA  475
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     704  AGGCTGCCATCTTTGCCAAGCACAGAGGTCGCCGGGAGAGCGGTTCGTCGTGTGGGGGCA  763

Qy     476  TAGTCATCAGCTCCTGCTGGATTCTCGCCGATTTCTCTGCTGGATCCTGCTGAGTCGTGA  535
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     764  TAGTCATCAGCTCCTGCTGGATTCTCGCCGATTTCTCTGCTGGATCCTGCTGAGTCGTGA  823

Qy     536  CAGTCACCGACAACATGCTGTGTGCTGGAGACTCGGAGCGGCGGCCCCCAGGCAAACT  595
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     824  CAGTCACCGACAACATGCTGTGTGCTGGAGACTCGGAGCGGCGGCCCCCAGGCAAACT  883
```

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   884 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCCGGTCGGTG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   944 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCCGGTCGGTG 1003

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1063

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1123

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1183

Qy   896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGCGGGCGCCCCAGGCAAACT 1243

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1303

Qy   1016 TGACTTTGGTGGGGCATCATCAGCTGGGGGCCTGGGCTGTGGGGGCTGGGCTGTGGGGGCTGGGGTG 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 TGACTTTGGTGGGGCATCATCAGCTGGGGGCCTGGGCTGTGGGGGCTGGGCTGTGGGGGCTGGGGTG 1363

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1416

RESULT 7
AR059996
LOCUS   AR059996   1419 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 62 from patent US 5840533.
ACCESSION   AR059996
VERSION   AR059996.1 GI:5986446
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1419)
  AUTHORS   Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
  TITLE   Tissue plasminogen activator
  JOURNAL   Patent: US 5840533-A 62 24-NOV-1998;
FEATURES   Location/Qualifiers
     source   1..1419
                     /organism="unknown"
BASE COUNT   309 a   416 c   423 g   271 t
ORIGIN

Query Match   94.4%; Score 1065; DB 6; Length 1419;
Best Local Similarity   99.5%;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   56 CCCAGGCCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   344 CCCCTGCCTCGTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 403

Qy   116 GCACGCACAGCCTCACCGAGCTGGGTGCCTCCTGGGTCCTGCCCTCTCCCGTGCCATGTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   404 GCACGCACAGCCTCACCGAGCTGGGTGCCTCCTGGGTCCTGCCCTCTCCCGTGCCATGTCCATGATCCTGA 463

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
REFERENCE   1 (bases 1 to 1419)
AUTHORS     Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
            Kobayashi,M.
TITLE       NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
JOURNAL     Patent: JP 1989104167-A 11 21-APR-1989;
            FUJISAWA PHARMACEUT CO LTD
COMMENT
OS  Artificial gene
OC  Artificial sequence; Genes.
PN  JP 1989104167-A/11
PD  21-APR-1989
PF  01-AUG-1988 JP 1988192320
PR  03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
    13-NOV-1987 GB 87 8726683
PI  NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
PI  NOTANI JOJI,
PI  KOBAYASHI MASAKAZU
PC  C12N9/64,C12N1/20,C12N5/00,C12N15/00//A61K37/54,(C12N9/64, PC
    C12R1:19),
PC  (C12N9/64,C12R1:91);
CC  strandedness: Double;
CC  topology: Linear;
CC  hypothetical: No;
CC  anti-sense: No;
CC  *source: clone=pmTQk112;
FH  Key             Location/Qualifiers
FT  CDS             1..1419
FT                  /product='new t-PA'.
FEATURES            Location/Qualifiers
     source         1..1419
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
BASE COUNT    309 a    413 c    426 g    271 t
ORIGIN

Query Match             94.4%;  Score 1065;  DB 6;  Length 1419;
Best Local Similarity   99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG 115
         |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   344 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG 403

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   404 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA 463

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   464 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 523

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGCCACGTGCCACGTGCTGAAGAACCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   524 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGCCACGTGCCACGTGCTGAAGAACCGCA 583

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGCCTGAGACAGTACA 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   584 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGCCTGAGACAGTACA 643

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
         ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   644 GCCAGCCCTCAGTTTCGCATCAAAGGAGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 703

Qy   416 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGAGCGGTTCCTGTGCGGGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   704 AGGCTGCCATCTTGCCAAGCACAGGAGGTCGCCGGAGCGGTTCCTGTGCGGGGGCA 763

Qy   476 TACTCATCAGCTCCTGCTGAATTCTCTCTGCGCCCCACTGCTTCCGC 535
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   764 TACTCATCAGCTCCTGCTGAATTCTCTCTGCGCCCCACTGCTTCCGC 823

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTGTCCCTGGGACGGAGAGC 595
```

```
Db   824 CCCACCACCTGACGGTGATCTTGGGCCAGAGACATACAGGTGCCTGGGCTGGTGCTCCTGGGCGAGGAGGAGC 883

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   884 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy   656 ATGACATTGCGCTGCGAGCTGAAATCGGATTCCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   944 ATGACATTGCGCTGCGAGCTGAAATCGGATTCCGTCCCGCTGTGCCCAGGAGAGCAGCG 1003

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1004 TGGTCCGCACTGTGTGCCTTCCCCGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1063

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGGTCTCCTTCATTCGGAGCGGCTGAAGG 835
         ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  1064 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGGTCTCCTTTCATTCGGAGCGGCTGAAGG 1123

Qy   836 AGGCTCATGTGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
         ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  1124 AGGCTCATGTGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1183

Qy   896 CAGTCACCGACAACATGTGTGTCTGGAGACACTCGGAGCGGCGGGCCCCAGCCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  1184 CAGTCACCGACAACATGTGTGTCTGGAGACACTCGGAGCGGCGGCGGCCCCAGCCAAACT 1243

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGAGCCCCCTGGTGTCTGAACGATGGCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1244 TGCACGACGCCTGCCAGGGCGATTCGGAGGAGCCCCCTGGTGTCTGAACGATGGCCGCA 1303

Qy  1016 TGACTTTGGTGGTGGGCCATCATCAGCTGGGCGTGGGCTGGTGACAGAAGAGATGTCCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1304 TGACTTTGGTGGTGGGCCATCATCAGCTGGGCGTGGGCTGGTGACAGAAGAGATGTCCCGGGTG 1363

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1416
```

```
RESULT 9
I06609
LOCUS       I06609                  1689 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 44 from Patent EP 0293934.
ACCESSION   I06609
VERSION     I06609.1  GI:590639
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1689)
AUTHORS     Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
            Hashimoto,A. and Yuzuriha,T.
TITLE       Mutant t-PA with kringle replacement
JOURNAL     Patent: EP 0293934-A1 44 07-DEC-1988;
FEATURES            Location/Qualifiers
     source         1..1689
                    /organism="unknown"
BASE COUNT    377 a    483 c    505 g    324 t
ORIGIN

Query Match             94.4%;  Score 1065;  DB 6;  Length 1689;
Best Local Similarity   99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG 115
         |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    56 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGTCAGCCTACCGTG 614

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA 673

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGGAATTCCATGATCCTGA 733
```

```
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1689)
  AUTHORS   Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S.,
            Hashimoto,A. and Yuzuriha,T.
  TITLE     Mutant t-PA with kringle replacement
  JOURNAL   Patent: EP 0293934-A1 46 07-DEC-1988;
FEATURES    Location/Qualifiers
  source    1..1689
            /organism="unknown"
BASE COUNT    377 a    484 c    504 g    324 t
ORIGIN

Query Match        94.4%;  Score 1065;  DB 6;  Length 1689;
Best Local Similarity 99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

Qy    56 CCCAGGCCGGCCCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  115
         ||||  ||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   614 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  673

Qy   116 GCACGACAGCCTCACCGAGTCGGGTGCTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA  175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 GCACGACAGCCTCACCGAGTCGGGTGCTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA  733

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   734 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  793

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCCAAGCCCTGGTGTGCCACGTGCTGAAGAACCGCA  295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   794 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCCAAGCCCTGGTGTGCCACGTGCTGAAGAACCGCA  853

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCTCCCGTGCTCTGAGACAGTACA  355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   854 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCTCCCGTGCTCTGAGACAGTACA  913

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGC  415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   914 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGC  973

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCGCCCACTGCCGTTTCCGC  475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGGTTCGCCCACTGCCGTTTCCGC  1033

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTGCTGCTGACCTGCAGCTGGACGCGGAGGTGTG  535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1034 TACTCATCAGCTCCTGCTGGATTCTCTGCTGCTGACCTGCAGCTGGACGCGGAGGTGTG  1093

Qy   536 CCCACCACCTGACGGTGATCTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC  595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1094 CCCACCACCTGACGGTGATCTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC  1153

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1213

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGGAGACTGGAGGCG  715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1214 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGGAGACTGGAGGCG  1273

Qy   716 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCGGAGACTGGACGGAGTGTG  775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1274 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCGGAGACTGGACGGAGTGTG  1333

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG  835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAGG  1393

Qy   836 AGGCTCATGTCAGACTGTACCAGCTGTGTGCTGGAGACACTCCGGAGCCGCGCACATCCAGCCGCG  895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1394 AGGCTCATGTCAGACTGTACCAGCTGTGTGCTGGAGACACTCCGGAGCCGCGCACATCCAGCCGCG  1453

Qy   896 CAGTCACCGACAACATGTCGTGTGCTGGAGACACTCGGAGCCGCGGCGGCCCCAGGCAAACT  955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1454 CAGTCACCGACAACATGTCGTGTGCTGGAGACACTCGGAGCCGCGGCGGCCCCAGGCAAACT  1513

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCTGGTGTGTCTGAACGATGGCCGCA  1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1514 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCTGGTGTGTCTGAACGATGGCCGCA  1573

Qy  1016 TGACTTTCGGTGGGCATCATCAGCTGCTGAGACAACTCGGACGAGGAGGATGTCCCGGGGTG  1075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1574 TGACTTTCGGTGGGCATCATCAGCTGCTGAGACAACTCGGACGAGGAGGATGTCCCGGGGTG  1633

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1634 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1686
```

RESULT 10
```
I06614
LOCUS       I06614          1689 bp    DNA       linear    PAT 02-DEC-1994
DEFINITION  Sequence 46 from Patent EP 0293934.
ACCESSION   I06614
VERSION     I06614.1  GI:590640
KEYWORDS    .
SOURCE      Unknown.
```

```
Qy    896  CAGTCACCGACAACATGCTGTGTGCTCGAGACACTCGGAGCGCGGCCCCAGGCAAACT  955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454  CAGTCACCGACAACATGCTGTGTGCTCGAGACACTCGGAGCGCGGCCCCAGGCAAACT 1513

Qy    956  TGCACGACGGCCTGCCAGGGCGATTCGGAGGAGGCCCCCTTGTGTCTGAACGATGGCCCGA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514  TGCACGACGGCCTGCCAGGGCGATTCGGAGGAGGCCCCCTTGTGTCTGAACGATGGCCGCA 1573

Qy   1016  TGACTTTGTGGGCATCATCAGCTGGGCGTGTGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||
Db   1574  TGACTTTGTGGGCATCATCAGCTGGGCGTGTGACAGAAGGATGTCCCGGGTG 1633

Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1634  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1686
```

RESULT 11
I08789

```
LOCUS       I08789               1689 bp    DNA    linear    PAT 02-DEC-1994
DEFINITION  Sequence 3 from Patent WO 8804690.
ACCESSION   I08789
VERSION     I08789.1  GI:588494
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1689)
  AUTHORS   Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.
  JOURNAL   Patent: WO 8804690-A 3 30-JUN-1988;
FEATURES             Location/Qualifiers
     source          1..1689
                     /organism="unknown"
BASE COUNT      379 a    480 c    505 g    325 t
ORIGIN
```

```
Query Match      94.4%;  Score 1065;  DB 6;  Length 1689;
Best Local Similarity 99.5%;  Pred. No. 6.5e-244;
Matches 1068;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

Qy     56  CCCAGGCGGCCCTCGAGGGAAACAGTGACAGTGACTCTACTTTGGGAATGGGTCAGCCTACCGTG  115
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614  CCCCTGCCTGCTCTGAGGGAAACAGTGACTGACTGACTTTGGGAATGGGTCAGCCTACCGTG  673

Qy    116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA  175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    674  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA  733

Qy    176  TAGGCAAGGTTTACACAGCACAGAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA  235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    734  TAGGCAAGGTTTACACAGCACAGAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA  793

Qy    236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCGCA  295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    794  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCGCA  853

Qy    296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCTGAGACAGTACA  355
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    854  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCTGAGACAGTACA  913

Qy    356  GCCAGCTCAGTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTCCCACCCCTGGC  415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    914  GCCAGCTCAGTTCGCATCAAAGGAGGGCTTCTTCGCCGACATCGCCTCCCACCCCTGGC  973

Qy    416  AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGAGGAGCCGGTTCGTGCGGGGCA  475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974  AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGAGGAGCCGGTTCGTGCGGGGCA 1033

Qy    476  TACTCATCAGCTCCGCTGGAATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC  535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1034  TACTCATCAGCTCCGCTGGAATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 1093

Qy    536  CCCACCACCTGACCGGTGATCTTGGGCAGAACATACGCGGTGGTCCCTGCGGGGAGAGC  595
```

```
Db   1094  CCCACCACCTGACCGGTGATCTTGGGCAGAGACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1153
           |||||||||||||||||||||||||||||||||||||||||

Qy    596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1154  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1213

Qy    656  ATGACATTGCGCTGCTGCAGCTCGAAATCGGATTCGTCCCGCTGTGCCAGGAGAGCCAGCG  715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1214  ATGACATTGCGCTGCTGCAGCTCGAAATCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCG 1273

Qy    716  TGGTCCGCACTGTGTGCCTTCCCCGGGCGACCTGCAGCTGCCCGGACTGCAGCGGACGAGTGTG  775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274  TGGTCCGCACTGTGTGCCTTCCCCGGGCGACCTGCAGCTGCCCGGACTGCAGCGGACGAGTGTG 1333

Qy    776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCGGAGCGGCGGCTGAAGG  835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCGGAGCGGCGGCTGAAGG 1393

Qy    836  AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCGAGACACTCGGAGCACATTACTTACAGAA  895
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1394  AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCGAGACACTCGGAGCACATTACTTAACAGAA 1453

Qy    896  CAGTGCCACCGACAACATGCTGGGGCGGGGGCCCCCAGGCAAACT  955
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1454  CAGTCACCGACAACATGCTGGGAGGCGGGGGCCCCCAGGCAAACT 1513

Qy    956  TGCACGACGGCCTGCCAGGGCGATTCGGAGGAGGCCCCCTGTGTCTGAACGATGGCCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514  TGCACGACGGCCTGCCAGGGCGATTCGGAGGAGGCCCCCTGTGTCTGAACGATGGCCGCA 1573

Qy   1016  TGACTTTGTGGGCATCATCAGCTGGGCGTGTGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1574  TGACTTTGTGGGCATCATCAGCTGGGCGTGTGACAGAAGGATGTCCCGGGTG 1633

Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1634  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1686
```

RESULT 12
A06611

```
LOCUS       A06611               1780 bp    DNA    linear    PAT 23-JUN-1993
DEFINITION  Synthetic nucleotide sequence of the 1.78kb t-PA BamHI fragment of
            pt-PABam.
ACCESSION   A06611
VERSION     A06611.1  GI:412906
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences
REFERENCE   1  (bases 1 to 1780)
  AUTHORS   .
  JOURNAL   Patent: WO 9000600-A 11 25-JAN-1990;
FEATURES             Location/Qualifiers
     source          1..1780
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS             17..1705
                     /codon_start=1
                     /transl_table=11
                     /product="t-PA analogue"
                     /protein_id="CAA00577.1"
                     /db_xref="GI:412907"
                     /translation="MDAMKRGLCCVLLCGAVFVSPSQEIHARPRRGARSYQVICRDE
                     KTQMIYQQHQSWLRPVLRSNRVEYCWCNSGRAQCHSVPVKSCSEPRCFNGGTCQQALY
                     PSDFVCQCPEGFAGKCCEIDTRATCYEDQGISYRGTWSTAESGAECTNWNSSALAQKP
                     YSGRRPDAIRLGLGNHNYCRNPDRDSKPWCYVFKAGKYSSEFCSTPACSEGNSDCYFG
                     NGSAYRGTHSLTESGASCLPWNSMILGKVYTAQNPSAQALGLGKHNYCRNPDGDAXP
                     WCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRR
                     SPGERPLCGGILISSCWILSAAHCFQERFPPHLLTVILGRTYRVVPGEEEQKPEVEKY
                     IVHKEFDDDTYDNIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYG
```

KHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA
CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP"

misc_difference   371
                   /note="'t' can be replaced by 'a'."
misc_difference   1367..1369
                   /note="'aag' can be replaced by 'agt' or 'tcc'."
misc_difference   1370..1372
                   /note="'cat' can be replaced by 'act'."

BASE COUNT   405 a   513 c   523 g   339 t
ORIGIN

Query Match   94.4%;   Score 1065; DB 6; Length 1780;
Best Local Similarity 99.5%;   Pred. No. 6.5e-244;
Matches 1068;   Conservative 0;   Mismatches 5;   Indels 0;   Gaps 0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   630 CCCTGCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 689

Qy   116 GCACGCACAGCCTCACCGACCTCACCGACCTCGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   690 GCACGCACAGCCCTCACCGACCTCACCGACCTCGGTGCCTCCTGCCTCCCGTGAATTCCATGATCCTGA 749

Qy   176 TAGGCAAGGTTTACACAGCACAGACACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   750 TAGGCAAGGTTTACACAGCACAGACACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 809

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCA 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCA 869

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTGCTCCACCTGCGCGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTGCTCCACCTGCGCGCCTGAGACAGTACA 929

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGCTCTTCGCCGACGACGCGTTCCTGTGCGGGGGCA 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGCTCTTCGCCGACGACGCGTTCCTGTGCGGGGGCA 989

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTGCGCCCGAGAGCGGTTCCTGCGCCGCCGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 AGGCTGCCATCTTTGCCAAGCACAGGAGTGCGCCCGAGAGCGGTTCCTGCGCCGCCGGCA 1049

Qy   476 TACTCATCAGCTCCTGCTGGAATTCTCTCTGCGCGCCCACTGCTTCTGCCGCCCACTGCT 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 TACTCATCAGCTCCTGCTGGAATTCTCTCTGCGCGCCCACTGCTTCTGCCGCCCACTGCT 1109

Qy   536 CCCACCACTGCACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGCAGGAGAGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 CCCACCACTGCACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGGCAGGAGAGC 1169

Qy   596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1229

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCGTGCCCGGAGAGCAGCG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGCGTGCCCGGAGAGCAGCG 1289

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGGACCTGCAGCCTGCGCCGGAACTGGACGGAGTGTG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 TGGTCCGCACTGTGTGCCTTCCCCGGGACCTGCAGCCTGCGCCGGAACTGGACGGAGTGTG 1349

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGCGCTTTCCTTTGTCTCCTTTCGGAGCCGTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 AGCTCTCCGGCTACGGCAAGCATGAGCGCTTTCCTTTGTCTCCTTTCGGAGCCGTGAAGG 1409

Qy   836 AGCGCTCATGTGCATCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410 AGCGCTCATGTGCATCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1469

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1470 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCCAGGCAAACT 1529

Qy   956 TGCACGAGGCCTGCCAGGCGATTCGGAGGAGGCCCCCTGTGTGTCTGAACGATGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 TGCACGAGGCCTGCCAGGCGATTCGGAGGAGGCCCCCTGTGTGTCTGAACGATGCCGCA 1589

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGGCCTGGGCTGTGGGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590 TGACTTTGGTGGGCATCATCAGCTGGGGCGGCCTGGGCTGTGGGACAGAAGGATGTCCCGGGTG 1649

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1702

RESULT 13
A10226
LOCUS   A10226   1780 bp   DNA   linear   PAT 17-FEB-1997
DEFINITION   Synthetic DNA for mutant tissue plasminogen activator (t-PA).
ACCESSION   A10226
VERSION   A10226.1 GI:490657
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1 (bases 1 to 1780)
  AUTHORS   Petersen,L.C. and Boel,E.
  TITLE   A tissue plasminogen activator analogue
  JOURNAL   Patent: EP 0351246-A 7 17-JAN-1990;
               NOVO-NORDISK A/S; NOVO NORDISK A/S
FEATURES   Location/Qualifiers
     source   1..1780
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS   17..1705
                     /codon_start=1
                     /transl_table=11
                     /product="mutant tissue plasminogen activator (t-PA)"
                     /protein_id="CAA00883.1"
                     /db_xref="GI:490658"
                     /translation="MDAMKRGLCVLLLCGAVFVFPSQEIHARFRRGARSYQVICRDE
                     KTQMIYQQHQSWLRPVLRSNRVEYCWCNSGRAQCHSVPVKSCSEPRCFNGGTCQQALY
                     FSDFVCQCPEGFAGKCCEIDTRATCYEDQGISYRGTWSTAESGAECTNWNSSALAQKP
                     YSGRRPDAIRLGLGNHNYCRNPDRDSKPWCYVFKAGKYSSEFCSTPACSEGNSDCYFG
                     NGSAYRGTHSLTESGASCLPWNSMILIGKVYTAQNPSAQALGLGKHNYCRNPDGDAKP
                     WCHVLKNRRLTWEYCDVPSCSTCGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRR
                     SPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRTYRVVPGEEEQKFEVEKY
                     IVHKEFDDDTYDNDIALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTECELSGYG
                     KHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA
                     CQGDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP"
     misc_difference 371..373
                     /note="in t-PA C87S,K419S"
                     /replace="agt"
     misc_difference 1367..1372
                     /note="in t-PA K419S, H420T"
                     /replace="agtact"
     misc_difference 1367..1369
                     /note="in t-PA C87S, K419S"
                     /replace="tcc"

BASE COUNT   405 a   513 c   523 g   339 t
ORIGIN

Query Match   94.4%;   Score 1065; DB 6; Length 1780;
Best Local Similarity 99.5%;   Pred. No. 6.5e-244;
Matches 1068;   Conservative 0;   Mismatches 5;   Indels 0;   Gaps 0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   630 CCCTGCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 689

Qy   116 GCACGCACAGCTTCACCGACCTCACCGAGAGACTCGGAGCGGCGGCCCCCGTGGAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
KEYWORDS      JP 198710457-A/1.
SOURCE        synthetic construct
  ORGANISM    synthetic construct
              artificial sequences.
REFERENCE     1  (bases 1 to 1836)
  AUTHORS     Pooru,P.H., Narendaa,K.K. and Shiyouguangu,R.R.
  TITLE       POLYCLINGLE PLASMINOGEN ACTIVATING FACTOR
  JOURNAL     Patent: JP 198710457-A 1 15-MAY-1987;
              AMERICAN HOME PROD CORP
COMMENT       OS  Artificial gene
              OC  Artificial sequence; Genes.
              OS  homo sapiens
              PN  JP 198710457-A/1
              PD  15-MAY-1987
              PF  13-AUG-1986 JP 198619173
              PR  14-AUG-1985 US 85   766163, 11-JUL-1986 US 86   884835 PI
              POORU POOUEN HANGU, NARENDAA KUMAARU KARUYAN, PI   SHIYOUGUANGU RIN
              RII
              PC  C12N9/52,A61K35/12,A61K35/74,A61K37/02,A61K37/54,C07H21/04, PC
              C12N15/00,
              PC  (C12N9/52,C12R1:19),(C12N15/00,C12R1:19);
              CC  strandedness: Double;
              CC  topology: Linear;
              CC  hypothetical: No;
              CC  anti-sense: No;
              CC  *source: cell_type=fibroblast;
              CC  *source: cell_line=WI-38cell;
              FH  Key             Location/Qualifiers
              FH
              FT  CDS             12..1793
              FT                  /product='this-kringle plasminogen activator'
              FT  sig_peptide     12..71
              FT                  /product='signal peptide of urokinase' FT
              mat_peptide     72..1790
              FT                  /product='this-kringle plasminogen activator'.
FEATURES      Location/Qualifiers
  source      1..1836
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
BASE COUNT    422 a    537 c    528 g    349 t
ORIGIN
```

```
Query Match              94.4%;    Score 1065;  DB 6;   Length 1836;
Best Local Similarity    99.5%;    Pred. No. 6.5e-244;
Matches 1068;   Conservative   0;   Mismatches   5;   Indels   0;  Gaps   0;
```

```
Qy    56 CCCAGGCGGCCTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   718 CCCCTGCCTGCTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 777

Qy   116 GCACGCAAGCCTCACCGAGTCGGGTGCCTCCTCGTCCCGTCGGAATTCCATGATCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 GCACGCAAGCCTCACCGAGTCGGGTGCCTCCTCGTCCCGTCGGAATTCCATGATCTGA 837

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGCCTGGGCACTGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGCCTGGGCACTGGCAAACATA 897

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCCAGCCCTGGTGCCAGTGCTGAAGAACCGGA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCCAGCCCTGGTGCCAGTGCTGAAGAACCGGA 957

Qy   296 GGCTGACGTGGGAGTACTGTGGATGTGCCCTCCTTGCCACTCGGCACGTGCTTGGGTGTGG 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GGCTGACGTGGGAGTACTGTGGATGTGCCCTCCTTGCCACTCGGCACGTGCTTGGGTGTGG 1017

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTTCGCCGACATCGCCTCTTCGCCACCCCTGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1018 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTTCGCCGACATCGCCTCTTCGCCACCCCTGC 1077

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGGTCGCCCGAGGAGGTCGGTTCCTGTGCGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   690 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCGTCCCGTCGGAATTCCATGATCTGA 749

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGCCTGGGCACTGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   750 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGCCTGGGCACTGGCAAACATA 809

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCCAGCCCTGGTGCCAGTGCTGAAGAACCCGA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCCAGCCCTGGTGCCAGTGCTGAAGAACCCGA 869

Qy   296 GGCTGACGTGGGAGTACTGTGGATGTGCCCTCCTTGCCACTCGGCACGTGCTTGGGTGTGG 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 GGCTGACGTGGGAGTACTGTGGATGTGCCCTCCTTGCCACTCGGCACGTGCTTGGGTGTGG 929

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTTCGCCGACATCGCCTCCCACCCCTGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTTCGCCGACATCGCCTCCCACCCCTGC 989

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGCGCAGGGAGCCGGTTCCTGTGCGGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 AGGCTGCCATCTTTGCCAAGCACAGGAGCGCAGGGAGCCGGTTCCTGTGCGGGGGCA 1049

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCACTGCTTCCAGGAGAGAGGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1050 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCACTGCTTCCAGGAGAGAGGTTTCCGC 1109

Qy   536 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1110 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1169

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1170 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1229

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGCCCAGGAGCAGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1230 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCGCCCAGGAGCAGCG 1289

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGGACTTGCAGCTGCGCCGGACTGGACGGAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1290 TGGTCCGCACTGTGTGCCTTCCCCCGGGACTTGCAGCTGCGCCGGACTGGACGGAGTGTG 1349

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGGCGGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1350 AGCTCTCCGGCTACGGCAAGCATGAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGGCGGCTGAAGG 1409

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1410 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1469

Qy   896 CAGTCACCGACAACACATGCTGTGTCTGAGACACTCGGAGACGGCGGCCCCAGGCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1470 CAGTCACCGACAACACATGCTGTGTCTGAGACACTCGGAGACGGCGGCCCCAGGCAAACT 1529

Qy   956 TGCACGACGCCTGCCAGGCGATTCGGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCGA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1530 TGCACGACGCCTGCCAGGCGATTCGGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCGA 1589

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGGTG 1075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1590 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGTGGACAGAAGGATGTCCCGGGGTG 1649

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1650 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1702
```

```
RESULT 14
E01176
LOCUS       E01176         1836 bp   DNA       linear      PAT 29-SEP-1997
DEFINITION  DNA encoding tris-kringle plasminogen activator.
ACCESSION   E01176
VERSION     E01176.1   GI:2169435
```

```
Db  1078 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGTTCCTGTGCGGGGGCA 1137
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   718 CCCCTGCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG  777

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGC   535
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCTCCGTGCCTCCGTGAATTCCATGATCCTGA  175

Db  1138 TACTCATCAGCTCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGC  1197
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCTCCGTGCCTCCGTGAATTCCATGATCCTGA  837

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTGCCTGGCGAGGAGAGC   595
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGCCCAGTGCCTGGGCACTGGGCAAACATA  235

Db  1198 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTGCCTGGCGAGGAGAGC  1257
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGCCCAGTGCCTGGGCACTGGGCAAACATA  897

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA   655
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGCGAAGAACCGCA  295

Db  1258 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGCGAAGAACCGCA  957

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCTTCCGGCTGGACGTGC   715
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGTGATGTGCCCTCCTGCGGCCTGACAGTACA  355

Db  1318 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCTTCCGGCTGGACGTGC  1377
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGTGATGTGCCCTCCTGCGGCCTGACAGTACA  1017

Qy   716 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCCGGGACTGGACGGAGTGTG   775
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  415

Db  1378 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCCGGGACTGGACGGAGTGTG  1437
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1018 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC  1077

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG   835
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGTTCCTGTGCGGGGGCA  475

Db  1438 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG  1497
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1078 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCGTTCCTGTGCGGGGGCA  1137

Qy   836 AGGCTCATGTGCAGCTGTACCATCCCAGCCGCTGCACATCACAACATTACTTAACAGAA   895
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   476 TACTCATCAGCTCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGC  535

Db  1498 AGGCTCATGTGCAGCTGTACCATCCCAGCCGCTGCACATCACAACATTACTTAACAGAA  1557
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1138 TACTCATCAGCTCCTGCTGGATTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGC  1197

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACT   955
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTGCCTGGCGAGGAGAGC  595

Db  1558 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACT  1617
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1198 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTGCCTGGCGAGGAGAGC  1257

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA  1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655

Db  1618 TGCACGACGCCTGCCAGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCCCA  1677
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1258 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  1317

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGTGGGACGAAGGAATGGTGTCCCGGGTG  1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCTTCCGGCTGGACGTGC  715

Db  1678 TGACTTTGGTGGGCATCATCAGCTGGGGCGTGGGACAGAAGGAATGGTGTCCCGGGTG  1737
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1318 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCTTCCGGCTGGACGTGC  1377

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG  1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   716 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCCGGGACTGGACGGAGTGTG  775

Db  1738 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG  1790
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1378 TGGTCCGGACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCCGGGACTGGACGGAGTGTG  1437

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG  835
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1438 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG  1497

Qy   836 AGGCTCATGTGCAGCTGTACCATCCCAGCCGCTGCACATCACAACATTACTTAACAGAA  895
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1498 AGGCTCATGTGCAGCTGTACCATCCCAGCCGCTGCACATCACAACATTACTTAACAGAA  1557

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACT  955
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1558 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGGCCCCAGGCAAACT  1617

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA  1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618 TGCACGACGCCTGCCAGGGCGATTCGGGAGGAGGCCCCCTGGTGTGTCTGAACGATGGCCCCA  1677

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGTGGGACGAAGGAATGGTGTCCCGGGTG  1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678 TGACTTTGGTGGGCATCATCAGCTGGGGCGTGGGACAGAAGGAATGGTGTCCCGGGTG  1737

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG  1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1738 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG  1790
```

RESULT 15

```
I01581
LOCUS       I01581       1836 bp ss-DNA       linear   PAT 21-MAY-1993
DEFINITION  Sequence 1 from Patent US 4916071.
ACCESSION   I01581
VERSION     I01581.1  GI:271057
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1836)
  AUTHORS   Hung,P.P., Kalyan,N.K. and Lee,S.-g.L.
  TITLE     Poly-kringle plasminogen activator
  JOURNAL   Patent: US 4916071-A 1 10-APR-1990;
            American Home Products Corporation; New York, NY
FEATURES             Location/Qualifiers
     source          1..1836
                     /organism="unknown"
BASE COUNT      422 a    537 c    528 g    349 t
ORIGIN
```

```
Query Match         94.4%;  Score 1065;  DB 6;  Length 1836;
Best Local Similarity   99.5%;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy   56 CCCAGGCGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
```

Search completed: August 19, 2003, 04:46:27

us-09-987-455-2.rge

Wed Aug 20 09:29:37 2003

Job time : 4739.31 secs

```
OM nucleic - nucleic search, using sw model

            Copyright  GenCore version 5.1.6
            Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:        August 18, 2003, 19:51:53 ; Search time 397.386 Seconds
                                            (without alignments)
                              7662.485 Million cell updates/sec

Title:          US-09-987-455-2
Perfect score:  1128
Sequence:       1 atgaaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters:      5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_19Jun03:*
   1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
   2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
   3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
   4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
   5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
   6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
   7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
   8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
   9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
  10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
  11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
  12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
  13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
  14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
  15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
  16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
  17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
  18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
  19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
  20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
  21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
  22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
  23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
  24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
  25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1128	100.0	1128	24	AAD40614	OmpA-K2S fusion pr
2	1065	94.4	1314	10	AAN91123	Sequence of coding
3	1065	94.4	1341	11	AAQ06762	Sequence encoding
4	1065	94.4	1419	10	AAN91133	Sequence of coding
5	1065	94.4	1689	9	AAN82178	Tissue plasminogen
6	1065	94.4	1780	11	AAQ01358	Sequence encoding
7	1065	94.4	2100	10	AAN91119	Sequence of native
8	1065	94.4	2100	11	AAQ05532	Plasmid pST112 enc
9	1065	94.4	2162	16	AAQ87370	Human tPA clone.
10	1065	94.4	2162	19	AAV37294	Human tissue plasm
11	1065	94.4	2547	7	AAN60659	Sequence encoding
12	1065	94.4	2560	10	AAN90542	Plasmid pKG12 cont
13	1065	94.4	7533	8	AAN70491	Entire sequence of
14	1065	94.4	7533	9	AAN81970	Plasmid pEMp1-tPA.
15	1064	94.3	1680	17	AAT27588	Novel plasminogen
16	1064	94.3	1680	17	AAT27587	Novel plasminogen
17	1063.4	94.3	1614	11	AAQ03581	Glycosylated plasm
18	1063.4	94.3	1614	12	AAQ11550	Sequence encoding
19	1063.4	94.3	1659	12	AAQ11551	Sequence encoding
20	1063.4	94.3	1686	21	ABL52893	Tissue type plasmi
21	1063.4	94.3	1689	9	AAN82179	Tissue plasminogen
22	1063.4	94.3	1689	11	AAQ04903	Mutated t-PA analo
23	1063.4	94.3	1689	11	AAQ04904	Part of tPA024 gen
24	1063.4	94.3	1689	11	AAQ05177	Part of tPA023 gen
25	1063.4	94.3	1689	12	AAQ13379	Sequence encoding
26	1063.4	94.3	1689	24	ABK52546	T-PA67+ mutant wit
27	1063.4	94.3	1740	17	AAT27585	Human tissue plasm
28	1063.4	94.3	1955	21	AAZ31998	Full-length tissue
29	1063.4	94.3	1989	7	AAN60700	Human tissue plasm
30	1063.4	94.3	2091	12	AAQ12073	Sequence encoding
31	1063.4	94.3	2099	11	AAQ05534	T-PA variant contg
32	1063.4	94.3	2100	12	AAQ12072	Plasmid plgNQTPA
33	1063.4	94.3	2170	8	AAN70248	T-PA with -ve char
34	1063.4	94.3	2170	16	AAQ86576	Human tissue plasm
35	1063.4	94.3	2249	11	AAQ05535	Plasmid plgN delta
36	1063.4	94.3	2313	11	AAQ05944	Sequence encoding
37	1063.4	94.3	2457	7	AAN60146	cDNA sequence enco
38	1063.4	94.3	2509	10	AAN90508	cDNA of human tiss
39	1063.4	94.3	2519	24	ABK52545	Human tissue plasm
40	1063.4	94.3	2519	24	ABL59544	Human tissue type
41	1063.4	94.3	2547	4	AAN30001	Sequence encoding
42	1063.4	94.3	7360	17	AAT15930	DHFR/intron (WTras
43	1062	94.1	1065	24	AAD40616	Human kringle 2 se
44	1062	94.1	1068	10	AAN91120	Sequence of coding

ALIGNMENTS

```
RESULT 1
AAD40614
ID   AAD40614 standard; DNA; 1128 BP.
XX
AC   AAD40614;
XX
DT   30-OCT-2002  (first entry)
XX
DE   OmpA-K2S fusion protein encoding DNA.
XX
KW   Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW   K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW   artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW   cerebroprotective; cardiant; ompA; fusion protein; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT
FT   CDS             1..1128
FT                   /*tag= a
FT                   /product= "OmpA-K2S fusion protein"
FT                   /transl_except= (pos:1126..1128, aa:Pro-Gly)
FT                   /note="No stop codon"
FT                   /partial
XX
PN   WO200240650-A2.
XX
PD   23-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-EP12857.
XX
```

```
PR   14-NOV-2000; 2000GB-0027779.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PI   Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;
XX
DR   WPI; 2002-519376/55.
DR   P-PSDB; AAE25034.
XX
PT   Producing active, correctly folded recombinant tissue plasminogen
PT   activator, Kringle 2 serine protease in prokaryotic cells by expressing
PT   the protein-encoding DNA operably linked to DNA coding for signal
PT   peptide OmpA -
XX
PS   Claim 7; Page 30; 80pp; English.
XX
CC   The present invention relates to a method of producing extracellularly
CC   secreted, active, correctly folded, recombinant tissue plasminogen
CC   activator (tPA), Kringle 2 serine protease molecule (K2S) or their
CC   variants in prokaryotic cells by expressing the protein-encoding DNA
CC   operably linked to DNA coding for signal peptide OmpA. The method is
CC   useful for producing recombinant DNA-derived tPA, K2S or their variants.
CC   Sequences of the invention are useful for manufacturing a medicament
CC   for treating stroke, cardiac infarction, acute myocardial infarction,
CC   pulmonary embolism, any artery occlusion such as intracranial artery
CC   occlusion (e.g., arteries supplying the brain), peripherally occluded
CC   arteries, coronary artery occlusion, deep vein thrombosis or related
CC   diseases associated with unwanted blood clotting. The present sequence
CC   is a DNA encoding a fusion protein comprising OmpA and K2S protein.
XX
SQ   Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;

Query Match             100.0%;  Score 1128;  DB 24;  Length 1128;
Best Local Similarity   100.0%;  Pred. No. 5.1e-239;
Matches  1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGAAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTCGCTACCGTGGCCCAG 60

Qy        61 GCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCAGCCTACCGTGGCACG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTCAGCCTACCGTGGCACG 120

Qy       121 CACAGGCCTCACCCAGTCGGGTGCCCTCCTCGCCTCCTGCGGTTCGGTTCGGTTCGGTGATAGGC 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CACAGGCCTCACCCAGTCGGGTGCCCTCCTCGCCTCCTGCGGTTCGGTTCGGTTCGGTGATAGGC 180

Qy       181 AAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCCTGGGCAAACATAATTAC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCCTGGGCAAACATAATTAC 240

Qy       241 TGCCGGAATCCTGATGGGGATGCCAAGCCTGTGCCACGTGGTGAAGAACCGCAGGCTG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TGCCGGAATCCTGATGGGGATGCCAAGCCTGTGCCACGTGGTGAAGAACCGCAGGCTG 300

Qy       301 ACGTGGGAGTACTGTGATGTGCCCTCCTGGCCTCCACCTGCGGCCTGAGCAGTACAGCCAG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ACGTGGGAGTACTGTGATGTGCCCTCCTGGCCTCCACCTGCGGCCTGAGCAGTACAGCCAG 360

Qy       361 CCTCAGTTTGCGATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTTGGCAGGCT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CCTCAGTTTGCGATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTTGGCAGGCT 420

Qy       421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTC 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTC 480

Qy       481 ATCAGCTCCTGCTGGATTCTCTCGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCAC 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ATCAGCTCCTGCTGGATTCTCTCGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCAC 540

Qy       541 CACCTGACGGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGGAGGAGCAGAAA 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CACCTGACGGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGGAGGAGCAGAAA 600

Qy       601 TTTGAAGTCGAAAAATACATTGTCCATAGGAATTCGATGATGACACTTACGACAATGAC 660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TTTGAAGTCGAAAAATACATTGTCCATAGGAATTCGATGATGACACTTACGACAATGAC 660

Qy       661 ATTGCGCTGCTGCGACGTGAAATCGGATTCGTCCCGCTGTCCGGTGCCCTGTGCCCGTGGTC 720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 ATTGCGCTGCTGCGACGTGAAATCGGATTCGTCCCGCTGTCCGGTGCCCTGTGCCCGTGGTC 720

Qy       721 CGGACTGTGTGCCTTCCCCCGGCGCCAGCTGCAGCCGACTGACGGAGTGTGAGCTC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CGGACTGTGTGCCTTCCCCCGGCGCCAGCTGCAGCCGACTGACGGAGTGTGAGCTC 780

Qy       781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGCT 840

Qy       841 CATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACACATTACTTAACAGAACAGTC 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACACATTACTTAACAGAACAGTC 900

Qy       901 ACCCACAACATGCTGTGTCGAGACACTCGGAGCGAGGCGGCGGGGCCCCCAGGCAAACTTGCAC 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 ACCGACAACATGCTGTGTCGAGACACTCGGAGCGAGGCGGCGGGGCCCCCAGGCAAACTTGCAC 960

Qy       961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACT 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCATGACT 1020

Qy      1021 TTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTTGGACAGAAGGATGTCCCGGGTGTGTAC 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 TTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTTGGACAGAAGGATGTCCCGGGTGTGTAC 1080

Qy      1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

```
RESULT 2
AAN91123
ID   AAN91123 standard; DNA; 1314 BP.
XX
AC   AAN91123;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pTQKrPA delta trp.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1065
FT                   /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988; 88EP-0112569.
XX
PR   03-AUG-1987; 87GB-0018298.
PR   26-OCT-1987; 87GB-0025052.
PR   13-NOV-1987; 87GB-0026681.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
```

PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94409.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1314 BP; 286 A; 386 C; 393 G; 249 T; 0 other;

Query Match   94.4%;   Score 1065; DB 10; Length 1314;
Best Local Similarity 99.5%;   Pred. No. 3.9e-225;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 298

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCTCCCGTGGAATTCCATGATCCTGA 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GCACGCACAGCCCTCACCGAGTCGGGTGCCTCCTGCCTCCTCCCGTGGAATTCCATGATCCTGA 358

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 418

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCCGCA 478

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACA 538

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 598

Qy   416 AGGCTGCCATCTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 AGGCTGCCATCTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCA 658

Qy   476 TACTCATCAGCTCCTGCTGGAATTCTCTCTGCCGCCCACTGCTTCTCCAGGAGAGGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TACTCATCAGCTCCTGCTGGAATTCTCTCTGCCGCCCACTGCTTCTCCAGGAGAGGTTTCCGC 718

Qy   536 CCCACCACCTGACGGTGATCTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   719 CCCACCACCTGACGGTGATCTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGAGC 778

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   779 AGAAATTTGAAGTCGCGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 838

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   839 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCGTCCCGCTGTGCCCAGGAGAGCAGCG 898

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCAGCTGGACCGGAGTGTG 775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   899 TGGTCCGCACTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCAGCTGGACCGGAGTGTG 958

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   959 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGG 1018

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1019 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1078
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAAGACACTCGGAGCGGCGGGCCCCAGGCCAAACT 955

Db 1079 CAGTCACCGACAACATGCTGTGTGCTGGAAGACACTCGGAGCGGCGGGCCCCAGGCCAAACT 1138
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   956 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1015

Db 1139 TGCACGACGGCCTGCCAGGGCGATTCGGGAGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1198
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1016 TGACTTTGTGGGCATCATCAGTGTGGGCTGGGCTGGTGCTGCAGAAGGATGTCCCGGGTG 1075

Db 1199 TGACTTTGTGGGCATCATCAGTGTGGGCTGGGCTGGTGCTGCAGAAGGATGTCCCGGGTG 1258
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1076 TGTACACAAAGGTTACCAACTACACTAGACTGGATTCGTGACAACATGGCACCG 1128

Db 1259 TGTACACAAAGGTTACCAACTACACTAGACTACCTAGGATTCGTGACAACATGGCACCG 1311

RESULT 3
AAQ06762
ID   AAQ06762 standard; DNA; 1341 BP.
XX
AC   AAQ06762;
XX
DT   09-JAN-2003 (updated)
DT   01-MAR-1991 (first entry)
XX
DE   Sequence encoding non-glycosylated tPA deriv.
XX
KW   Tissue plasminogen activator; thrombolysis; finger domain;
KW   EGF domain; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   4..1341
FT   /*tag= a
FT   /product=tPA deriv.
XX
PN   EP400545-A.
XX
PD   05-DEC-1990.
XX
PF   28-MAY-1990;   90EP-0110096.
XX
PR   14-JUL-1989;   89DE-3923339.
PR   31-MAY-1989;   89DE-3917781.
XX
PA   (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI   Stern A, Kohnert U, Rudolph R, Fischer S, Martin U;
XX
DR   WPI; 1990-363094/49.
DR   P-PSDB; AAR08150.
XX
PT   New non-glycosylated form of tissue plasminogen activator - with
PT   thrombolytic activity and long plasma life
XX
PS   Claim 2; page 16; 21pp; German.
XX
CC   This sequence encodes a non-glycosylated tPA deriv. lacking the
CC   finger- and EGF-domains. The polypeptide has a lower clearance
CC   rate (longer half-life) than natural tPA while retaining thrombo-
CC   lytic activity and stimulation by fibrin.
CC   (Updated on 09-JAN-2003 to add missing OS field.)
XX
SQ   Sequence 1341 BP; 293 A; 397 C; 399 G; 252 T; 0 other;

Query Match   94.4%;   Score 1065; DB 11; Length 1341;
Best Local Similarity 99.5%;   Pred. No. 3.9e-225;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

```
RESULT 4
AAN91133
ID   AAN91133 standard; DNA; 1419 BP.
XX
AC   AAN91133;
XX
DT   25-MAR-2003 (updated)
DT   03-OCT-2002 (updated)
DT   18-JUN-1990 (first entry)
XX
DE   Sequence of coding region in plasmid pmTQk112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1065
FT                  /*tag=  a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;  88EP-0112569.
XX
PR   03-AUG-1987;  87GB-0018298.
PR   26-OCT-1987;  87GB-0025052.
PR   13-NOV-1987;  87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94416.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
PS   Disclosure; Page 7; 68pp; English.
XX
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;

Query Match       94.4%;  Score 1065;  DB 10;  Length 1419;
Best Local Similarity 99.5%;  Pred. No. 3.9e-225;
Matches 1068;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;
```

(sequence alignment of Qy and Db follows)

```
FT                              /product=modified tPA
XX
PN   JP63230083-A.
PD   26-SEP-1988.
XX
PF   20-MAR-1987;   87JP-0064339.
XX
PR   20-MAR-1987;   87JP-0064339.
XX
PA   (EISA ) EISAI CO LTD.
XX
DR   WPI; 1988-311961/44.
DR   P-PSDB; AAP82581.
XX
PT   Modified tissue plasminogen activator - having glycine-183 and serine-186
PT   residues sustd. with serine and threonine.
XX
PS   Disclosure; Page ?; 16pp; Japanese.
XX
CC   One N-glycosylation site, i.e. NSS (117-119) is substituted with
CC   NSM and the N-glycosylation is removed.
CC   Plasmid encoding the modified tPA is 99-6300 and its transformant is
CC   E.coli RR1-Zem 99-6300 (FERM P-9127).
CC   This modified tPA, used to treat thrombosis, is of high quality and
CC   has a longer half life period in blood.
CC   See also AAN2177-N82179.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

Query Match             94.4%;  Score 1065;  DB 9;  Length 1689;
Best Local Similarity   99.5%;  Pred. 4e-225;
Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy       56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      614 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 673

Qy      116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCTCCCGTGCCTCCTCCGTGAATTCCATGATCCTGA 175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      674 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCGTGCCTCCTCCCGTGCCTCCTCCGTGAATTCCATGATCCTGA 733

Qy      176 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCCAGGCACTGCCCAGGCACTGCCCAGGCCTGGGCCTGGGCAAACATA 235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      734 TAGGCAAGGTTTACACAGCACAGACCCCAGTGCCCAGGCACTGCCCAGGCACTGCCCAGGCCTGGGCCTGGGCAAACATA 793

Qy      236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      794 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 853

Qy      296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCCGGCGCCTGAGACAGTACA 355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      854 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCCGGCGCCTGAGACAGTACA 913

Qy      356 GCCAGCCTCAGTTTGCATCAAAGGAGGGCTCTTCGCCGACATGCGCTCCCACCCCTGGC 415
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      914 GCCAGCCTCAGTTTGCATCAAAGGAGGGCTCTTCGCCGACATGCGCTCCCACCCCTGGC 973

Qy      416 AGGCTGCCATCTTTGCCAAGCACAGGAGGCACCAGGAGGTCGCCGGCAGGCGGTTCCGC 475
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      974 AGGCTGCCATCTTTGCCAAGCACAGGAGGCACCAGGAGGTCGCCGGCAGGCGGTTCCGC 1033

Qy      476 TACTCATCAGCTCCTGCCTGGATTCGTGACAACATGCGCACCG 535
            |||||||||||||||||||||||||||||||||||||||||||
Db     1034 TACTCATCAGCTCCTGCCTGGATTCGTGACAACATGCGCACCG 1093

Qy      536 CCCACCACCTGACGCGTGATCTTGGGCAGAACATACCGGGTGTCCTGCGGAGGAGC 595
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1094 CCCACCACCTGACGCGTGATCTTGGGCAGAACATACCGGGTGTCCTGCGGAGGAGC 1153
```

And the left column (rotated, second alignment set):

```
Qy      356 GCCAGCCTCAGTTTGCATCAAAGGAGGGCTCTTCGCCGACATGCGCTCCCACCCCTGGC 415
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      644 GCCAGCCTCAGTTTGCATCAAAGGAGGGCTCTTCGCCGACATGCGCTCCCACCCCTGGC 703

Qy      416 AGGCTGCCATCTTTGCCAAGCACAGGAGGCACCAGGAGGTCGCCGGCAGGCGGTTCCGC 475
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      704 AGGCTGCCATCTTTGCCAAGCACAGGAGGCACCAGGAGGTCGCCGGCAGGCGGTTCCGC 763

Qy      476 TACTCATCAGCTCCTGCCTGGATTCTCTCGCCGCCACTGCTTCCAGGAGGTTTCCGC 535
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      764 TACTCATCAGCTCCTGCCTGGATTCTCTCGCCGCCACTGCTTCCAGGAGGTTTCCGC 823

Qy      536 CCCACCACCTGACGGGTGATCTTGGGCAGAACATACCGGGTGTCCTGCGGAGGAGC 595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      824 CCCACCACCTGACGGGTGATCTTGGGCAGAACATACCGGGTGTCCTGCGGAGGAGC 883

Qy      596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      884 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy      656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCG 715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      944 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCG 1003

Qy      716 TGGTCCGCACTGTGTGCCTTCCCCCGGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1004 TGGTCCGCACTGTGTGCCTTCCCCCGGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1063

Qy      776 AGCTTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG 835
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      836 AGCTTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG 835

Qy      836 AGGGCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG 895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1064 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG 1123

Qy      836 AGGGCTCATGTCAGACTGTCAGACCGTACCCATCCAGCCGCGTGCACATCACAACATTTACTTAACAGAA 895
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1124 AGGGCTCATGTCAGACTGTCAGACCGTACCCATCCAGCCGCGTGCACATCACAACATTTACTTAACAGAA 1183

Qy      896 CAGTCACCGACAACAATGCTGTGTGCTGCGAGACACTCGGAGGCGGCGGCCCCAGGCAAACT 955
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1184 CAGTCACCGACAACAATGCTGTGTGCTGCGAGACACTCGGAGGCGGCGGCCCCAGGCAAACT 1243

Qy      956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCCTCGTGTCTGAACGATGGCCGCA 1015
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1244 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGAACGATGGCCCCA 1303

Qy     1016 TGACTTTCGTGGGCATCATCAGCTGGGGCTGTGGACAGAAGGATGTCCCGGGTG 1075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1304 TGACTTTGGTGGGCATCATCAGCTGGGGCTGTGGACAGAAGGATGTCCCGGGTG 1363

Qy     1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACTGCGACCG 1128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1364 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACTGCGACCG 1416
```

```
RESULT 5
AAN82178
ID   AAN82178  standard; DNA; 1689 BP.
XX
AC   AAN82178;
XX
DT   25-MAR-2003  (updated)
DT   10-MAR-2003  (updated)
DT   03-NOV-1990  (first entry)
XX
DE   Tissue plasminogen activator with S-119 substd for M.
XX
KW   Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1689
FT                   /*tag= a
```

```
XX
DR   WPI; 1990-016567/03.
DR   P-PSDB; AAR04699.
XX
PT   New tissue plasminogen activator (t-PA) analogue -
PT   with higher fibrin selectivity than native t-pa, useful for
PT   treating infarction, thrombosis and embolism
XX
PS   Disclosure; Fig 5A-D; 24pp; English.
XX
CC   It is modified in the patent by replacing one or more codons specifying
CC   a positively charged amino acid.  The resulting analogues are inserted
CC   into a replicable expression vector which is used to transform or
CC   transfect a host cell which is grown to express a t-PA analogue.  The
CC   t-PA analogue is useful for the treatment of diseases or disorders
CC   associated with the formation of thrombi in blood vessels e.g.
CC   infarctions, thrombosis and embolism.  The analogue, in the 1-chain form,
CC   exhibits the properties of a proenzyme; however on plasmin-catalysed
CC   cleavage of the one-chain form, the activity of the 2-chain form is fully
CC   retained.  Compared to native t-PA, it has a higher fibrin selectivity,
CC   this results in a fibrinolytic agent with a higher fibrin selectivity,
CC   as the fibrinogenolytic activity induced by the analogue is reduced
CC   relative to that induced by native 1-chain t-PA.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;

Query Match             94.4%;  Score 1065;  DB 11;  Length 1780;
Best Local Similarity   99.5%;  Pred. No. 4e-225;
Matches 1068;  Conservative   0;  Mismatches   5;  Indels   0;  Gaps   0;

Qy    56 CCCAGGCGGCCTTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTG 115
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   630 CCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCTACCGTG 689

Qy   116 GCACGCAAGCCTCACCCAGTGACGTGCCGTGCCGTGGTGCCTCCTGCCGTGCTGAAGAACCGCA 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   690 GCACGCAAGCCTCACCCAGTGACGTGCCGTGCCGTGGTGCCTCCTGCCGTGCTGAAGAACCGCA 749

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAAGCCACTGGGCCTGGGCAAACATA 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   750 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAAGCCACTGGGCCTGGGCAAACATA 809

Qy  -236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 869

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCTCCTGCCTCCACCTGGGGCTGCGAGCAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 GGCTGACGTGGGAGTACTGTGATGTGCCTCCTGCCTCCACCTGGGGCTGCGAGCAGTACA 929

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGAGGGGCTCTTCGCCGACATGCCCACCCCTGGC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 GCCAGCCTCAGTTTCGCATCAAAGAGGAGGGGCTCTTCGCCGACATGCCCACCCCTGGC 989

Qy   416 AGGCTGCCATCTTGCCAAGCACAGAGAGTCGCCGGAGAGCCGTTCCTGTGCGGGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 AGGCTGCCATCTTGCCAAGCACAGAGAGTCGCCGGAGAGCCGTTCCTGTGCGGGGGCA 1049

Qy   476 TACTTCATCAGCTCCTGGATTCTCTGCTGGATTCTCTCCAGGAGGAGGTTTCCGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1050 TACTTCATCAGCTCCTGGATTCTCTGCTGGATTCTCTCCAGGAGGAGGTTTCCGC 1109

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1110 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1169

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1170 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1229

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCGCCAGGAGCAGCG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1213

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCGCCAGGAGCAGCG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1214 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCGCCAGGAGCAGCG 1273

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCCAGCTGGACGGAGTGTG 775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1274 TGGTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCCAGCTGGACGGAGTGTG 1333

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1334 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1393

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1394 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1453

Qy   896 CAGTCACCCACAACACATGCTGTGTGGAGGAGGCCCCCTGTGTCTGAACGATGGCCGCA 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1454 CAGTCACCCACAACACATGCTGTGTGGAGGAGGCCCCCTGTGTCTGAACGATGGCCGCA 1513

Qy   956 TGCACGACGCCCTGCCAGGGCGATTCGGGAGGCCCCTGCGTGTCTGAACGATGGCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1514 TGCACGACGCCCTGCCAGGGCGATTCGGGAGGCCCCTGCGTGTCTGAACGATGGCCGCA 1573

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGGGCCTGGGCGTGGGCGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1574 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGGGCCTGGGCGTGGGCGGTG 1633

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACAACATGGCGACCG 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1634 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTTGACAACATGGCGACCG 1686
```

```
RESULT 6
AAQ01358
ID   AAQ01358 standard; DNA; 1780 BP.
XX
AC   AAQ01358;
XX
DT   25-MAR-2003 (updated)
DT   08-SEP-1990 (first entry)
XX
DE   Sequence encoding wild type tissue plasminogen activator (t-PA).
XX
KW   Wild type tissue plasminogen activator (t-PA); infarction treatment;
KW   thrombosis treatment; embolism treatment.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   misc_feature   1..123
FT                  /*tag= a
FT                  /note="synthetic 5' adaptor"
FT   misc_feature   1738..1780
FT                  /*tag= b
FT                  /note="synthetic 3' adaptor"
XX
PN   EP351246-A.
XX
PD   17-JAN-1990.
XX
PF   14-JUL-1989;   89EP-0307194.
PR   15-JUL-1988;   88DK-0003952.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Petersen LC, Boel E;
```

```
PS   Example 29; Fig 21; 68pp; English.
XX
CC   cDNA sequence of native tPA gene is excised from plasmid pST112, and
CC   digested with BamHI and SalI to form plasmid pST18.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match             94.4%;  Score 1065;  DB 10;  Length 2100;
Best Local Similarity   99.5%;  Pred. No. 4.1e-225;
Matches 1068;   Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    56 CCCAGGCGCCTCTGAGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
         |||  | |   ||||||| ||||||||||||||||||||||||||||||||||||||
Db   638 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  697

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA  175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   698 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGAATTCCATGATCCTGA  757

Qy   176 TAGGCAAAGTTTACACAGCACAGAACCCAGTGCCCAGCACTGGGCTGGGCAAACATA  235
         |||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db   758 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGCACTGGGCTGGGCAAACATA  817

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   818 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  877

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   878 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  937

Qy   356 GCCAGCCTCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCCGCTCCCACCCTGGC  415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   938 GCCAGCCTCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCCGCTCCCACCCTGGC  997

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA  475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   998 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA  1057

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTGCGGTCTTCTCTGCCACGCCCACTGCTTCCGC  535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1058 TACTCATCAGCTCCTGCTGGATTCTCTGCGGTCTTCTCTGCCACGCCCACTGCTTCCGC  1117

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGC  595
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1118 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAGGAGC  1177

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACA  655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1178 AGAAATTTGAAGTCGAAAAATACATTGTCATAAGGAATTCGATGATGACACTTACGACA  1237

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCGTGTGCCCAGGAGCAGCG  715
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1238 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCGTGTGCCCAGGAGCAGCG  1297

Qy   716 TGGTCCGACTGTGTCCTTCCCCCGGCGGACCTGCCAGCTGCCGGACTGGAGTGTG  775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1298 TGGTCCGACTGTGTCCTTCCCCCGGCGGACCTGCCAGCTGCCGGACTGGAGTGTG  1357

Qy   776 AGTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1358 AGTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  1417

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACAACATTTACTTAACAGAA  895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1418 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACAACATTTACTTAACAGAA  1477

Qy   896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGACGCGGCGCCCCAGGCAAACT  955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1478 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGACGCGGCGCCCCAGGCAAACT  1537
```

```
Db  1230 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCGTGTGCCCAGGAGCAGCG  1289
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   716 TGGTCCGACTGTGTCCTTCCCCCGGCGGACCTGCCAGCTGCCGGACTGGAGTGTG  775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1290 TGGTCCGACTGTGTCCTTCCCCCGGCGGACCTGCCAGCTGCCGGACTGGAGTGTG  1349

Qy   776 AGTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1350 AGTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG  1409

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACAACATTTACTTAACAGAA  895
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1410 AGGCTCATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACAACATTTACTTAACAGAA  1469

Qy   896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCGCCCCAGGCAAACT  955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1470 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCGCCCCAGGCAAACT  1529

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGCCCCTGTGTTCTGAACGATGGCCGCA  1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1530 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGCCCCTGTGTTCTGAACGATGGCCGCA  1589

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGTGGACAGAAGGATGTGCCCGGGTG  1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1590 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGTGGACAGAAGGATGTGCCCGGGTG  1649

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGGCGACCG  1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1650 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACATGGCGACCG  1702
```

```
RESULT 7
AAN91119
ID   AAN91119 standard; DNA; 2100 BP.
XX
AC   AAN91119;
XX
DT   25-MAR-2003   (updated)
DT   03-OCT-2002   (updated)
DT   18-JUN-1990   (first entry)
XX
DE   Sequence of native tPA in plasmid pST112.
XX
KW   Tissue plasminogen activator; tPA; thrombolytic agent;
KW   plasminogen; vascular diseases.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            25..1710
FT                  /*tag= a
XX
PN   EP302456-A.
XX
PD   08-FEB-1989.
XX
PF   02-AUG-1988;   88EP-0112569.
XX
PR   03-AUG-1987;   87GB-0018298.
PR   26-OCT-1987;   87GB-0025052.
PR   13-NOV-1987;   87GB-0026683.
XX
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
DR   WPI; 1989-040625/06.
DR   P-PSDB; AAP94406.
XX
PT   New tissue plasminogen activator -
PT   comprising finger and growth factor domains lacking tPA for
PT   longer half-life and stronger thrombolytic activity.
XX
```

```
Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCGTGCCTCCCGTGGAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    698 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCGTGCCTCCCGTGGAATTCCATGATCCTGA 757

Qy    176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGCCTGGCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    758 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGCCTGGCAAACATA 817

Qy    236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    818 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 877

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCGTCTCGGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    878 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCGTCTCGGGCCTGAGACAGTACA 937

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    938 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 997

Qy    416 AGGCTGCCATCTTTGCCAAGCACCAGAGTCGCCCGGAGAGCGGTTCCTCTGTGCGGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    998 AGGCTGCCATCTTTGCCAAGCACCAGAGTCGCCCGGAGAGCGGTTCCTCTGTGCGGGGGCA 1057

Qy    476 TACTCATCAGCTCCTGCTGGATTCTCTGCGCCACTGCTTCCAGGAGAGGTTTCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1058 TACTCATCAGCTCCTGCTGGATTCTCTGCGCCACTGCTTCCAGGAGAGGTTTCCGC 1117

Qy    536 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1118 CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1177

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1178 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTACGACA 1237

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGCCGCTGTCCCGCTGTGCCCAGGAGAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1238 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGCCGCTGTCCCGCTGTGCCCAGGAGAGCAGCG 1297

Qy    716 TGGTCCGGACTGTGTGCCTTCCCCGGGACTGCAGCTGCGCCGACTGCGACGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1298 TGGTCCGGACTGTGTGCCTTCCCCGGGACTGCAGCTGCGCCGACTGCGACGGAGTGTG 1357

Qy    776 AGCCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGAGCGCGGCTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1358 AGCCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGAGCGCGGCTGAAGG 1417

Qy    836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1418 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1477

Qy    896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCACTCGGAGCACCGGCAAACT 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1478 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCACTCGGAGCACCGGCAAACT 1537

Qy    956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCGCCCCTGTCTGTCTGAACGATGCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1538 TGCACGACGCCTGCCAGGCGATTCGGGAGGCGCCCCTGTCTGTCTGAACGATGCCGCA 1597

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1598 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGCGTGTGGACAGAAGGATGTCCCGGGTG 1657

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1710
```

RESULT 9
AAQ87370
ID   AAQ87370 standard; DNA; 2162 BP.
XX

```
Qy     56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    638 CCCTGCCTGCTGTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 697
```

RESULT 8
AAQ05532
ID   AAQ05532 standard; DNA; 2100 BP.
XX
AC   AAQ05532;
XX
DT   25-MAR-2003 (updated)
DT   11-DEC-1990 (first entry)
XX
DE   Plasmid pSTl12 encoding novel N-terminal for tissue plasminogen
DE   activator (tPA).
XX
KW   Fibrin; clotting; thrombolytic; vascular disease; stroke;
KW   myocardial infarction; heart attack; pulmonary embolism; ds;
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   25..1710
FT   /*tag= a
FT   mat_peptide   130..1710
FT   /*tag= b
XX
PN   EP379890-A.
XX
PD   01-AUG-1990.
XX
PF   10-JAN-1990; 90EP-0100457.
XX
PR   23-JAN-1989; 89GB-0001422.
XX
PA   (FUJI) FUJISAWA PHARM CO LTD.
XX
PI   Niwa M, Satoh S, Suzuki S, Otsuka K, Kusunoki C;
XX
DR   WPI; 1990-232757/31.
DR   P-PSDB; AAR06237.
XX
PT   New tissue plasminogen activator - having N-terminal peptide of
PT   plasminogen linked to tissue plasminogen activator for increased
PT   stability in vivo.
XX
PS   Disclosure; Page ?; ?pp; English.
XX
CC   tPA with N-terminal peptide of plasminogen linked is more stable in
CC   vivo than the native form. It is useful as a thrombolytic agent in
CC   the treatment of vascular diseases eg myocardial infarction,
CC   pulmonary embolism etc.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match   94.4%; Score 1065; DB 11; Length 2100;
Best Local Similarity   99.5%; Pred. No. 4.1e-225;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCGCCCCTGTCTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1538 TGCACGACGCCTGCCAGGCGATTCGGGAGGCGCCCCTGTCTGTCTGAACGATGGCCGCA 1597

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1598 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGCGTGTGGACAGAAGGATGTCCCGGGTG 1657

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1710
```

```
AC   AAQ87370;
XX
DT   25-MAR-2003 (updated)
DT   19-SEP-1995 (first entry)
XX
DE   Human tPA clone.
XX
KW   Intron; recombination; combinatorial gene; trans-splicing;
KW   gene therapy; polymerase chain reaction; PCR; primer; amplification;
KW   tissue plasminogen activator; tPA; plasmid TPA-KS+; thrombolytic;
KW   ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     82..334
FT                   /*tag= a
FT                   /note="signal sequence and finger-like domain"
FT                   335..447
FT                   /*tag= b
FT                   /product= EGF-like domain
FT   CDS             448..714
FT                   /*tag= c
FT                   /product= Kringle-1 domain
FT   CDS             715..972
FT                   /*tag= d
FT                   /product= Kringle-2 domain
FT   CDS             973..2162
FT                   /*tag= e
FT                   /product= catalytic domain
XX
PN   WO9507351-A1.
XX
PD   16-MAR-1995.
XX
PF   12-SEP-1994; 94WO-US10146.
XX
PR   10-SEP-1993; 93US-0119512.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Jarrell KA;
XX
DR   WPI; 1995-123425/16.
XX
PT   New intron-mediated recombinant techniques - used for the
PT   generation and selection of novel genes and gene prods. for use
PT   in therapy
XX
PS   Example 4; Page 59-60; 87pp; English.
XX
CC   A cDNA clone of human tissue plasminogen activator (tPA) was
CC   amplified by PCR using the primers given in AAQ87368-69. The
CC   amplified tPA DNA (AAQ87370) was ligated into vector KS+ to
CC   obtain plasmid TPA-KS+. The construct was used in combinatorial
CC   methods involving RNA splicing-mediated shuffling of tPA domains,
CC   in plasmid pINVI (AAQ87347) to generate novel tPAs having
CC   improved thrombolytic properties.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;
```

```
Query Match           94.4%;  Score 1065;  DB 16;  Length 2162;
Best Local Similarity 99.5%;  Pred. No. 4.1e-225;
Matches 1068;  Conservative 0;  Mismatches 0;  Indels 5;  Gaps 0;

Qy     56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
           ||| || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    695 CCCCTGCCGCTGCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 754

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGGAATTCCATGATCCTGA 175
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    755 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCCGTGCCTCCCCGTGGAATTCCATGATCCTGA 814
```

```
Qy    176 TAGGCAAGGTTTACACAGACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    815 TAGGCAAGGTTTACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 874

Qy    236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    875 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 934

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    935 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 994

Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCCACCCCTGGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    995 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATGCCTCCCACCCCTGGC 1054

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGCGGTCGCCCGGAGCGGTTCCTGTGCGGGGGCA 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055 AGGCTGCCATCTTTGCCAAGCACAGGAGCGGTCGCCCGGAGCGGTTCCTGTGCGGGGGCA 1114

Qy    476 TACTCATCAGCTCCTGGATTCTCTGCGCCCCACTGCTTCCAGGAGAGGTTTCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115 TACTCATCAGCTCCTGGATTCTCTGCGCCCCACTGCTTCCAGGAGAGGTTTCCGC 1174

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCAGGAGGAGC 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCAGGAGGAGC 1234

Qy    596 AGAAATTTGAAGTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235 AGAAATTTGAAGTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294

Qy    656 ATGACATTGCGCTGCTGCTGAGCTGAAATCGGAATCGTCCCGCTGTGCCCGTGTGCCAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295 ATGACATTGCGCTGCTGCTGAGCTGAAATCGGAATCGTCCCGCTGTGCCCGTGTGCCAGCAGCG 1354

Qy    716 TGGTCCGACTGTGTGCCTTCCCCGGACTTGCAGTCGCAGCTGCCGGACTGGACGGAGTGTG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355 TGGTCCGACTGTGTGCCTTCCCCGGACTTGCAGTCGCAGCTGCCGGACTGGACGGAGTGTG 1414

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGCTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCGCTGAAGG 1474

Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1475 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 1534

Qy    896 CAGTGCACCGACAACATGCTGTGTCTGAGAACACTCGGAGCGGCGGGGCGGGCCCCAGGCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1535 CAGTGCACCGACAACATGCTGTGTCTGAGAACACTCGGAGCGGCGGGGCGGGCCCCAGGCAAACT 1594

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1595 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCGCA 1654

Qy   1016 TGACTTTGGTGGGGCATCATCAGCTGGGCGGCGTGGGACAGAAGGATGTCCCGGGTG 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1655 TGACTTTGGTGGGGCATCATCAGCTGGGCGGCGTGGGACAGAAGGATGTCCCGGGTG 1714

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1767
```

```
RESULT 10
AAV37294
ID   AAV37294 standard; DNA; 2162 BP.
XX
AC   AAV37294;
XX
DT   10-SEP-1998 (first entry)
```

XX
DE   Human tissue plasminogen activator gene sequence.
XX
KW   Plasmid pINVI; reverse-splicing intron; group II intron;
KW   exon binding site; domain V motif; branch site acceptor;
KW   nucleophilic group; transesterification; phosphodiester bond;
KW   autocatalytic Y-branched intron; reverse splicing reaction; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   82..334
FT   /tag= a
FT   misc_feature   335..447
FT   /note= "signal sequence and finger-like domain"
FT   misc_feature   448..714
FT   /tag= b
FT   /note= "EGF-like domain"
FT   misc_feature   715..972
FT   /tag= c
FT   /note= "Kringle-1 domain "
FT   misc_feature   973..2162
FT   /tag= d
FT   /note= "Kringle-2 domain"
FT   /tag= e
FT   /note= "catalytic domain"
XX
PN   US5780272-A.
XX
PD   14-JUL-1998.
XX
PF   07-JUN-1995;   95US-0488015.
XX
PR   10-SEP-1993;   93US-0119512.
XX
PA   (HARD) HARVARD COLLEGE.
XX
PI   Jarrell KA;
XX
DR   WPI; 1998-413060/35.
XX
PT   Reverse splicing construct containing fragments of autocatalytic
PT   introns - able to cleave and ligate discontinuous nucleic acid for
PT   generating new genes and e.g. ribozymes, libraries of enzymes and
PT   antibodies
XX
PS   Example 4; Columns 53-56; 56pp; English.
XX
CC   The present sequence represents the human tissue plasminogen activator
CC   gene. It was used to construct plamsid TPA-KS+, which is used in the
CC   course of the invention. The specification describes a purified
CC   reverse splicing intron which comprises a segment comprising a
CC   5'-part of a group II intron, including an exon binding site not
CC   naturally present in the intron and a second segment comprising a
CC   3'-part of a group II intron, including a domain V motif, a branch
CC   site acceptor, and a nucleophilic group for transesterifying a
CC   phosphodiester bond of an RNA. Together the two segments form an
CC   autocatalytic Y-branched intron which catalyses integration of at least
CC   the first segment into substrate RNA by a reverse splicing reaction.
CC   The reverse-splicing introns are used, by specific cleavage and ligation
CC   of discontinuous nucleic acid, to generate new genes and gene products,
CC   e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC   manipulation, e.g. replacements for restriction enzymes) or
CC   immunologically active e.g. signal-transducing proteins such as antibody
CC   and enzyme libraries.
XX
SQ   Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match   94.4%;   Score 1065;   DB 19;   Length 2162;
Best Local Similarity   99.5%;   Pred. No. 4.1e-225;
Matches 1068;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

Qy   56 CCCAGGCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115

RESULT 11

us-09-987-455-2.rng

```
AAN60659
ID   AAN60659 standard; DNA; 2547 BP.
XX
AC   AAN60659;
XX
DT   25-MAR-2003  (updated)
DT   22-JUL-1991  (first entry)
XX
DE   Sequence encoding human pre-tissue plasminogen activator (pre-t-PA).
XX
KW   Plasminogen conversion; vascular disease therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             87..191
FT                   /*tag= a
FT   mat_peptide     192..1775
FT                   /*tag= b
XX
PN   GB2173804-A.
XX
PD   22-OCT-1986.
XX
PF   21-APR-1986;   86GB-0609683.
XX
PR   22-APR-1985;   85US-0725468.
PR   01-APR-1986;   86US-0846697.
PR   22-APR-1985;   85US-0725468.
PR   01-APR-1986;   86US-0846697.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Heyneker HL, Vehar GA;
XX
DR   WPI; 1986-280715/43.
DR   P-PSDB; AAP60790.
XX
PT   New mutant forms of human tissue plasminogen activator- having
PT   higher specific activity and resistance to conversion to two chain
PT   form
XX
PS   Example; Fig 2; 34pp; English.
XX
CC   The patentors claim a novel recombinant human t-PA which is
CC   resistant to specific enzymatic cleavage because it is stabilised by
CC   site-directed mutagenesis at a 2-chain cleavage site. Partic. the
CC   natural Arg at position 275 is replaced by Gly or Glu, or Ile at
CC   position 276 is replaced (275 and 276 refer to the posn. of the AAs
CC   in the mature protein; i.e. AAs 310 and 311 of AAP60790). Also new are
CC   (1) DNA sequences; (2) expression vectors; and (3) microorganisms
CC   and cell cultures transformed with these vectors.
CC   (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;
```

Query Match   94.4%; Score 1065; DB 7; Length 2547;
Best Local Similarity 99.5%; Pred. No. 4.2e-225;
Matches 1068; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

```
Qy          56 CCCAGGCGGCCCTCGAGGGAAACAGTGACTGACTTTGGGAATGGGTCAGCCTACCGTG 115
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         700 CCCCTGCTGCTCTGAGGGAAACAGTGACTGACTTTGGGAATGGGTCAGCCTACCGTG 759

Qy         116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         760 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 819

Qy         176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         820 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 879

Qy         236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGCTGCCACGTGCCACGTGCTGAAGAACCCCA 295
```

RESULT 12
```
AAN90542
ID   AAN90542 standard; cDNA; 2560 BP.
XX
AC   AAN90542;
XX
DT   25-MAR-2003  (updated)
DT   05-JUN-1990  (first entry)
XX
DE   Plasmid pKG12 contg. DNA encoding human melanoma t-PA.
XX
KW   Human tissue plasminogen-activator gene; pKG12; ss.
XX
```

```
Db   880 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   296 GGCTGACGTGGGAGTACTGCAGGGATGTGCCTCCTGCTCCACCTGCGGCGGCCTGAGACAGTACA 355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   940 GGCTGACGTGGGAGTACTGCAGGGATGTGCCTCCTGCTCCACCTGCGGCGGCCTGAGACAGTACA 999

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1000 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 1059

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCA 475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1060 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCA 1119

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCGCTGCTTCCTGCTTCCTTCCAGGAGAGGTTTCCGC 535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1120 TACTCATCAGCTCCTGCTGGATTCTCTCGCTGCTTCCTGCTTCCTTCCAGGAGAGGTTTCCGC 1179

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1180 CCCACCACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1239

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1240 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1299

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCG 715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1300 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCG 1359

Qy   716 TGGTCCGACTGTGTGCCTTCCCCGGCGGACTTGCCAGCTGCCGACTGCCGACGGAGTGTG 775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1360 TGGTCCGACTGTGTGCCTTCCCCGGCGGACTTGCCAGCTGCCGACTGCCGACGGAGTGTG 1419

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1420 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1479

Qy   836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACATTACTTAACAGAA 895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1480 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACATTACTTAACAGAA 1539

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1540 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1599

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGCGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1600 TGCACGACGCCTGCCAGGGCGATTCGGAGGCGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1659

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGTGGACAGAAGGATGTCCCGGGTG 1075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1660 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGTGGACAGAAGGATGTCCCGGGTG 1719

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1720 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCCGACCG 1772
```

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     102..206
FT                   /*tag= a
FT   mat_peptide     ..1787
FT                   /*tag= b
FT                   /product=t-PA
XX
PN   EP297066-A.
XX
PD   28-DEC-1988.
XX
PF   14-JUN-1988;    88EP-0850207.
XX
PR   18-JUN-1987;    87SE-0002562.
XX
PA   (KABI ) KABIGEN AB.
XX
PI   Pohl G, Hansson L, Loewenadler B;
XX
DR   WPI; 1989-001503/01.
DR   P-PSDB; AAP93716.
XX
PT   Modified tissue plasminogen-activator - having domains deleted and amino
PT   acid changes to increase biological half-life and reduce inactivation.
XX
PS   Disclosure; Fig 1; 18pp; English.
XX
CC   The sequence was obtained from cDNA prepd. from mRNA extracted from
CC   Bowes melanoma cells. The cDNA was used to construct a gene library which
CC   was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated
CC   and shown to comprise the whole coding region for human t-PA as well as
CC   102bp 5'flanking, 760 bp 3' flanking DNA and a poly A tail.
CC   See also AAN91608.
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;
```

```
Query Match             94.4%;  Score 1065;  DB 10;  Length 2560;
Best Local Similarity   99.5%;  Pred. No. 4.2e-225;
Matches 1068;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy    56  CCCAGGCGCCGCCTCTGAGGGAAACAGTGAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG   115
              || ||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715  CCCCTGCCTCTGAGGGAAACAGTGAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG   774

Qy   116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA   175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   775  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA   834

Qy   176  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACAAACATA   235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACAAACATA   894

Qy   236  ATTACTGCCGGAATCCTGATGGGGATGGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA   295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895  ATTACTGCCGGAATCCTGATGGGGATGGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA   954

Qy   296  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGGCCTGAGACAGTACA   355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   955  GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGGCCTGAGACAGTACA   1014

Qy   356  GCCAGCCTCAGTTTCGCATCAAAGAGAGGCTCTTCGCCGACATCGCCTTCCCACCCCTGGC   415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1015  GCCAGCCTCAGTTTCGCATCAAAGAGAGGCTCTTCGCCGACATCGCCTTCCCACCCCTGGC   1074

Qy   416  AGGCTGCCATCTTTGCCAAGCACAGGAGGCGGTTCGCCGAGAGCGGTTCCTGTCGGGGGCA   475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1075  AGGCTGCCATCTTTGCCAAGCACAGGAGGCGGTTCGCCGAGAGCGGTTCCTGTCGGGGGCA   1134

Qy   476  TACTCATCAGCTCCTGCTGGATTCTCTCGCCCCACTGCTTCCGC
          ||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1135  TACTCATCAGCTCCTGCTGGATTCTCTCGCCCCACTGCTTCCGGAGAGGTTTCCGC   1194
                                                       |||||||||||||||
Qy   536  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTTCCCTGCGAGGAGGAGC   595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1195  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTTCCCTGCGAGGAGGAGC   1254

Qy   596  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA   655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255  AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA   1314

Qy   656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCGCCAGGAGCAGCG   715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCGCCAGGAGCAGCG   1374

Qy   716  TGGTCCGCACTGTGTGCCTTCCCCCGGCGACTGCAGCTGCCGGACTGGACCGAGTGTG   775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1375  TGGTCCGCACTGTGTGCCTTCCCCCGGCGACTGCAGCTGCCGGACTGGACCGAGTGTG   1434

Qy   776  AGCTCTCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG   835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435  AGCTCTCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG   1494

Qy   836  AGGCTCATGTCAGACTGTACCATCCAGCGCGTGCACATCACAACATTTACTTAACAGAA   895
          |||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db  1495  AGGCTCATGTCAGACTGTACCCATCCAGCCGCGTGCACATCACAACATTTACTTAACAGAA   1554

Qy   896  CAGTCACCGACAACATGCTGTGTGCTGAGACACTCCGAGGCGGCGGCCCCCAGGCAAACT   955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1555  CAGTCACCGACAACATGCTGTGTGCTGAGACACTCCGAGGCGGCGGCCCCCAGGCAAACT   1614

Qy   956  TGCACGACGCCTGCCAGGGACGGATTCGGAGGCGGATTCGGAGCGGTGTCTGAACGATGGCCGCA   1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1615  TGCACGACGCCTGCCAGGGACGGATTCGGAGGCGGATTCGGAGCGGTGTCTGAACGATGGCCGCA   1674

Qy  1016  TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCTGTGAGACAGAAGGATGTCCCGGGTG   1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1675  TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCTGTGAGACAGAAGGATGTCCCGGGTG   1734

Qy  1076  TGTACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG   1128
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1735  TGTACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG   1787
```

```
RESULT 13
AAN70491
ID   AAN70491 standard; cDNA; 7533 BP.
XX
AC   AAN70491;
XX
DT   25-MAR-2003 (updated)
DT   31-OCT-2002 (updated)
DT   07-JUN-1991 (first entry)
XX
DE   Entire sequence of high level expression vector pEMp1-tpa for tissue
DE   plasminogen activator (TPA).
XX
KW   Enhancer-blocking element; expression vector; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   mRNA            20..1708
FT                   /*tag= a
FT   misc_feature    7..1748
FT                   /*tag= b
FT                   /label= TPA cDNA
FT   3'UTR           1749..1955
FT                   /*tag= c
FT                   /note="from immunoglobulin kappa light chain gene"
FT   3'clip          1956..2164
FT                   /*tag= d
FT                   /note= "as above"
FT   misc_feature    2165..4463
```

```
FT                      /tag= e
FT                      /label= pBR322
FT                      /note= "contg. the ampicillinase gene and the
FT                      bacterial origin of replication"
FT     misc_feature     complement (4464..6436)
FT                      /tag= f
FT                      /label= E.coli gpt gene
FT     promoter         6437..6547
FT                      /tag= g
FT                      /label= SV40 origin region
FT     misc_signal      6454..6468
FT                      /tag= h
FT                      /label= SV40 (gpt) CAP site
FT     enhancer         6548..6715
FT                      /tag= i
FT                      /label= SV40 enhancer (TZ bp repeat)
FT     promoter         6716..6971
FT                      /tag= j
FT                      /label= lambda-1
FT                      /note= "functions as a blocking element"
FT     enhancer         6972..7281
FT                      /tag= k
FT                      /note= "from immunoglobulin heavy chain (IgH)"
FT     promoter         7282..7466
FT                      /tag= l
FT                      /note= "from metallothionine 1 (MT)"
FT     misc_signal      7467..7533
FT                      /tag= m
FT                      /label= MT CAP site
FT     polyA_signal     1930..1935
FT                      /tag= n
XX
PN   EP237157-A.
XX
PD   16-SEP-1987.
XX
PF   27-JAN-1987; 87EP-0300658.
XX
PR   07-MAR-1986; 86US-0837595.
XX
PA   (DAMO-) DAMON BIOTECH INC.
PA   (ABBO ) ABBOTT BIOTECH INC.
XX
PI   Gillies SD;
XX
DR   WPI; 1987-258421/37.
XX
PT   DNA construct for high level expression of protein - having a
PT   blocking element so that marker protein is produced only at
PT   levels required for selection
XX
PS   Disclosure; Fig. 9A-D; 20pp; English.
XX
CC   Vector pEMp1-tpa is an example of a vector of the invention. Pref.
CC   the blocking element is a promoter sequence oriented with its native
CC   3' end disposed proximal to the enhancer element and its 5' end
CC   disposed distal to the enhancer element. Alternatively the
CC   interposed blocking element comprises a promoter sequence with its 5'
CC   end disposed proximal to the enhancer element and, at its 3' end, a
CC   gene encoding for a different protein of interest. (see FT). The
CC   vector pEMp1 was constructed by inserting an enhancer-blocking
CC   element into pEM1.
CC   TPA cDNA fragment was inserted into the unique XhoI site present in
CC   both pEM1 and pEMp1. Of 26 transformants obtd. with pEM1-tpa, 5
CC   produced TPA at levels of 140-500 IU/ml. Of 16 transformants obtd.
CC   with pEMp1-tpa, 7 produced TPA at levels of 1000-6000 IU/ml.
CC   (Updated on 31-OCT-2002 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 7533 BP; 1921 A; 1899 C; 1844 G; 1868 T; 1 other;
```

```
Matches 1068;  Conservative  0;  Mismatches  0;  Indels  5;  Gaps  0;

Qy    56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
         |||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   633 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG   692

Qy   116 GCACGCACAGCCTCACCGGTGCGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA  175
         |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   693 GCACGCACAGCCTCACCGGTGCGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA  752

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGTGCCCAGGCACTGGGCCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   753 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGTGCCCAGGCACTGGGCCAAACATA 812

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   813 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA  872

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCACCTGCGGCCTGAGACAGTACA  355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   873 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA  932

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCCACATGCCTCCCACCCCTGGC  415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   933 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCCGCACATGCCTCCCACCCCTGGC  992

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA  475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   993 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1052

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCCGCCGCCACTGCTTCCAGGAGAGGTTCCGC  535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1053 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCCGCCGCCACTGCTTCCAGGAGAGGTTCCGC 1112

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGGAGC  595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1113 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGGAGC 1172

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1173 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1232

Qy   656 ATGACATTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGGCCGAGGAGCAGCG  715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1233 ATGACATTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGGCCGAGGAGCAGCG 1292

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACTGCAGCTGCCGACTGGACGGAGTGTG  775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1293 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACTGCAGCTGCCGACTGGACGGAGTGTG 1352

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGG  835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1353 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGG 1412

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCCACATCACAACATTACTTAACAGAA  895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1413 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCCACATCACAACATTACTTAACAGAA 1472

Qy   896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCGCCCAGGCAAACT  955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1473 CAGTCACCCACAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCGCCCCAGGCAAACT 1532

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCTGTGTGTCTGAACGATGCCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1533 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCTGTGTGTCTGAACGATGCCCGCA 1592

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTTGGACAGAAGGATGTCCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1593 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTTGGACAGAAGGATGTCCCGGGTG 1652

Qy  1076 TGTACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1653 TGTACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1705
```

```
Query Match      94.4%;  Score 1065;  DB 8;  Length 7533;
Best Local Similarity  99.5%;  Pred. No. 4.8e-225;
```

RESULT 14
AAN81970
ID   AAN81970 standard; DNA; 7533 BP.
XX
AC   AAN81970;
XX
DT   12-OCT-1990 (first entry)
XX
DE   Plasmid pEMp1-tPA.
XX
KW   Tissue plasminogen activator; myocardial infarction; thrombosis;
KW   embolism; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_RNA   7..1748
FT   /*tag= a
FT   /label=tPA_cDNA
FT   CDS   20..1708
FT   /*tag= b
FT   /product=t-PA
FT   3'UTR   1749..1955
FT   /*tag= c
FT   /label=K_3'UT
FT   /note="from Ig kappa light chain gene"
FT   polyA_signal   1930..1935
FT   /*tag= d
FT   misc_RNA   1956..2104
FT   /*tag= e
FT   /note="kappa 3 flanking sequence"
FT   misc_RNA   2105..4463
FT   /*tag= f
FT   /label=pBR322
FT   /note="contg. ampicillinase gene and origin of
FT   replication"
FT   misc_RNA   4464..6971
FT   /*tag= g
FT   /label=pSV2-gpt
FT   /note="non-coding strand"
FT   protein_bind   6437..6547
FT   /*tag= h
FT   /label=Cap_site
FT   /note="from SV40(gpt)"
FT   misc_RNA   6453..6469
FT   /*tag= i
FT   /label=origin-region
FT   /note="from SV40"
FT   repeat_unit   6548..6619
FT   /*tag= j
FT   /rpt_type=direct
FT   repeat_unit   6620..6691
FT   /*tag= k
FT   /rpt_type=direct
FT   enhancer   6633..6650
FT   /*tag= l
FT   misc_RNA   6716..6945
FT   /*tag= m
FT   /note="from SV40"
FT   protein_bind   6946..6971
FT   /*tag= n
FT   /label=blocking element
FT   /note="from lambda promoter"
FT   enhancer   6972..7281
FT   /*tag= o
FT   /label=cap_site
FT   /note="lambda mRNA"
FT   promoter   7282..7466
FT   /*tag= p
FT   /note="from IgH"
FT   protein_bind   7467..7533
FT   /*tag= q
FT   /label=cap_site
FT   /note="from metallothioneine I"
XX
PN   WO8800242-A.
XX
PD   14-JAN-1988.
XX
PF   25-JUN-1987; 87WO-US01569.
XX
PR   26-JUN-1986; 86US-0879038.
XX
PA   (DAMO-) DAMON BIOTECH INC.
XX
PI   Gillies SD;
XX
DR   WPI; 1988-021592/03.
DR   P-PSDB; AAP81913.
XX
PT   Prodn. of human tissue plasminogen activator - by myeloma cells,
PT   pref. in presence of epsilon amino caproic acid.
XX
PS   Disclosure; Page ?; p; English.
XX
CC   The plasmid was prepd by cleaving the very long 3' UT of t-PA cDNA
CC   (which causes mRNA instability) 34 nucleotides downstream of stop
CC   codon and inserting it intoa pEMp1 vector. The pEMp1 vector was
CC   constructed from the following components (see feature table):
CC   (a) a 2.25 PvuII-BamHI fragment from pSV2-gpt contg. the SV40
CC   enhancer and early region promoter, the E.coli gpt gene, the SV40
CC   small tumour antigen intervening sequence, and the SV40 termination
CC   and polyadenylation signals;
CC   (b) a 2.3 kb PvuII-EcoRI fragment from pBR322 contg. the
CC   ampicillinase gene and the bacterial origin of replication;
CC   (c) a 0.3 kb PvuII-EcoRI fragment contg. an Ig heavy chain enhancer;
CC   (d) a 0.25 kb SacI-BglII fragment contg. the metallothioneine I romoter;
CC   (e) a 0.4 kb AvaII-HaeIII fragment from the 3' UT of Ig kappa light
CC   chain gene; and
CC   (f) a 0.26 kb XbaI-bstNI fragment contg. the Ig light chain promoter,
CC   TATAA sequence, and the initiation site.
CC   The expression plasmid can be used to transform host cells esp.
CC   J5581, ATCC CRL 9132.
XX
SQ   Sequence 7533 BP; 1918 A; 1892 C; 1847 G; 1875 T; 1 other;

Query Match   94.4%;   Score 1065;   DB 9;   Length 7533;
Best Local Similarity   99.5%;   Pred. No. 4.8e-225;
Matches 1068; Conservative   0; Mismatches   5;   Indels   0;   Gaps   0;

Qy   56 CCCAGGCGGCCCTCTGAGGGAAACAGTGACTACTTTGGGAATGGGTCAGCCTACCGTG 115
              |||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Db   633 CCCCTGCCTGCTTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 692

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   693 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGA 752

Qy   176 TAGGCAAGTTTACACAGCACAGAACCCCAGTGCCGAGCACTGGGCCTGGGCACAAACATA 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   753 TAGGCAAGTTTACACAGCACAGAACCCCAGTGCCGAGCACTGGGCCTGGGCACAAACATA 812

Qy   236 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGA 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   813 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGA 872

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCTCCACCTGCCGCCTGCGCGCGCCTGAGACAGTACA 355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   873 GGCTGACGTGGGAGTACTGTGATGTGCCTCCACCTGCCGCCTGCGCGCGCCTGAGACAGTACA 932

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTTCCACCCCTGGC 415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   933 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCACATCCCCCTGGC 992

PN   US5504001-A.
XX
PD   02-APR-1996.
XX
PF   06-JUN-1994;   94US-0254485.
XX
PR   25-NOV-1987;   87US-0125629.
PR   28-JAN-1992;   92US-0827587.
PR   06-JUN-1994;   94US-0254485.
XX
PA   (ZYMO) ZYMOGENETICS INC.
XX
PI   Foster DC;
XX
DR   WPI; 1996-187699/19.
DR   P-PSDB; AAR96223.
XX
PT   Hybrid plasminogen activator comprises human tPA activator and
PT   N-terminal crosslinking domain from alpha2-plasmin inhibitor -
PT   useful to treat thrombosis and image blood clots
XX
PS   Example 3; Fig 11; 35pp; English.
XX
CC   A DNA construct (AAT27586) codes for a novel plasminogen
CC   activator (AAR96222) in which the kringle K1 domain (AAR96221)
CC   of plasminogen replaces the native K1 domain of tissue
CC   plasminogen activator (tPA) (AAR96220). It was obtd. by
CC   inserting a synthetic plasminogen K1 domain DNA (AAT27586)
CC   into tPA cDNA and can be expressed in transformed Escherichia
CC   coli RRI cells (FERM P-9272). Novel plasminogen activators are
CC   produced that show increased clot lysing specificity or plasma
CC   half-life.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 1680 BP; 387 A; 471 C; 479 G; 343 T; 0 other;

Query Match   94.3%; Score 1064; DB 17; Length 1680;
Best Local Similarity   99.5%; Pred. No. 6.6e-225;
Matches 1067; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   57 CCAGGCGGCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGG 116
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 CCTGGAATGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGG 665

Qy   117 CACGGACCAGCCTCACCGAGTCGGGTGCCTCCTCCCGTGAATTCCATGATCCTGAT 176
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 CACGGACCAGCCTCACCGAGTCGGGTGCCTCCTCCCGTGAATTCCATGATCCTGAT 725

Qy   177 AGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCTTGGGCAAAACATAA 236
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 AGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCTTGGGCAAAACATAA 785

Qy   237 TTACTGCCGAATCCTGATGGGGATGCCAAGCCCTGGTGTCCACGTGCTGAAGAACCGCAG 296
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 TTACTGCCGAATCCTGATGGGGATGCCAAGCCCTGGTGTCCACGTGCTGAAGAACCGCAG 845

Qy   297 GCTGACGTGGGAGTACTGTGATGTGCCCTGCCCCTGAGACAGTACAG 356
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   846 GCTGACGTGGGAGTACTGTGATGTGCCCTGCCCCTGAGACAGTACAG 905

Qy   357 CCAGCCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCA 416
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 CCAGCCCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCA 965

Qy   417 GGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCAT 476
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 GGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCAT 1025

Qy   477 ACTCATCAGCTCCTGCTGATTCTCTCCGCCCCACTGCTTCCAGCCGGTTTCCGCC 536
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 ACTCATCAGCTCCTGCTGATTCTCTCCGCCCCACTGCTTCCAGCCGGTTTCCGCC 1085

Qy   537 CCACCACCTGACCGTGATCTTGGGCAGAACATACCGGGTCCCGGCGGTGGAGAGCA 596

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   993 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGAGCGGTTCCTGTGCGGGGGCA 1052

Qy   476 TACTCATCAGCTCCTGGATTCTCTCGCCACTGCTTCCGCCGACTGCTTCCAGGAGAGAGTTCCGC 535
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1053 TACTCATCAGCTCCTGGATTCTCTCGCCACTGCTTCCGCCGACTGCTTCCAGGAGAGAGTTCCGC 1112

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1113 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGC 1172

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGCCATAAGGAATTCGATGATGACACTTACGACA 655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1173 AGAAATTTGAAGTCGAAAAATACATTGCCATAAGGAATTCGATGATGACACTTACGACA 1232

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCG 715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1233 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCG 1292

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACCTGCCAGCTGCCAGCTGGACGGAGTGTG 775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1293 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGACCTGCCAGCTGCCAGCTGGACGGAGTGTG 1352

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1353 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1412

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAA 895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1413 AGGCTCATGTCAGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAA 1472

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1473 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1532

Qy   956 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCGATTCGGGAGGAGGCCCCTTGGTGTCTGAACGATGGCCCA 1015
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1533 TGCACGACGCCTGCCAGGCGATTCGGGAGGCCGATTCGGGAGGAGGCCCCTTGGTGTCTGAACGATGGCCCA 1592

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGCTGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1593 TGACTTTGGTGGGCATCATCAGCTGGGCTGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1652

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128

Db   1653 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1705

RESULT 15
AAT27588
ID   AAT27588 standard; DNA; 1680 BP.
XX
AC   AAT27588;
XX
DT   25-MAR-2003 (updated)
DT   06-AUG-1996 (first entry)
XX
DE   Novel plasminogen activator DNA.
XX
KW   Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
KW   fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
KW   protein engineering; kringle; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   sig_peptide   1..105
FT   /*tag= a
FT   mat_peptide   106..1677
FT   /*tag= b
XX

```
Db  1086 CCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGGTCCCTGCGCGAGGAGGAGCA 1145
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   597 GAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAA  656
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1146 GAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAA 1205

Qy   657 TGACATTGCGCTGCTGACGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGT  716
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1206 TGACATTGCGCTGCTGCGACGCTGAAATCGGATTCGTCCCCGCTGTGCCCAGGAGCAGCGT 1265

Qy   717 GGTCCGCGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCCGAGACTGGACGGAGTGTGA  776
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1266 GGTCCGGACACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCCGACTGGACGGAGTGTGA 1325

Qy   777 GCTCTCCGGCTACCGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGA  836
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1326 GCTCTCCCGGCTACCGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGA 1385

Qy   837 GGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAAC  896
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1386 GGCTCATGTCAGACTGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAAC 1445

Qy   897 AGTCACCGACAACATGCTGTGTGCTGGAGCACTCGGAGCGGCGGCCCCAGGCAAACTT  956
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1446 AGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGCCCCAGGCAAACTT 1505

Qy   957 GCACGACGCCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCCGCAT 1016
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1506 GCACGACGCCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGGCCCGCAT 1565

Qy  1017 GACTTTGGTGGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGT 1076
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1566 GACTTTGGTGGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTGT 1625

Qy  1077 GTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
           ||||||||||||||||||||||||||||||||||||||||||
Db  1626 GTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1677
```

Search completed: August 18, 2003, 23:28:48
Job time : 400.386 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 88.2804 Seconds
                                         (without alignments)
                                         5639.757 Million cell updates/sec

Title: US-09-987-455-2
Perfect score: 1128
Sequence: 1 atgaaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 569978 segs, 22069156 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents NA:*
             1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
             2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
             3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
             4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
             5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
             6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Match Length	DB	ID	Description
1	1065	94.4	1419	2	US-08-811-949-62	Sequence 62, Appl
2	1065	94.4	1848	2	US-08-814-412-10	Sequence 10, Appl
3	1065	94.4	2101	2	US-08-811-949-42	Sequence 42, Appl
4	1065	94.4	2162	1	US-08-119-512-3	Sequence 3, Appli
5	1065	94.4	2162	1	US-08-488-015B-3	Sequence 3, Appli
6	1065	94.4	2162	1	US-08-488-015B-25	Sequence 25, Appl
7	1063.4	94.3	1314	2	US-08-811-949-48	Sequence 48, Appl
8	1063.4	94.3	1738	6	5200340-1	Patent No. 5200340
9	1063.4	94.3	1955	2	US-08-883-795A-39	Sequence 39, Appl
10	1063.4	94.3	2457	6	5344773-1	Patent No. 5344773
11	1063.4	94.3	7360	6	US-08-286-740-1	Sequence 1, Appli
12	1063.4	94.3	7360	5	PCT-US95-09576-1	Sequence 1, Appli
13	1062	94.1	1068	2	US-08-811-949-44	Sequence 44, Appl
14	1061.8	94.1	1314	2	US-08-811-949-50	Sequence 50, Appl
15	1061.8	94.1	1974	2	US-08-811-949-38	Sequence 38, Appl
16	1060.4	94.0	1170	2	US-08-811-949-64	Sequence 64, Appl
17	1058.8	93.9	1068	2	US-08-811-949-46	Sequence 46, Appl
18	1057	93.7	1314	2	US-08-811-949-54	Sequence 54, Appl
19	1056.4	93.7	1068	1	US-08-137-116-2	Sequence 2, Appli
20	1054.8	93.5	1065	1	US-08-427-640-1	Sequence 1, Appli
21	1054.8	93.5	1137	4	US-09-553-498-9	Sequence 9, Appli
22	1054.8	93.5	1137	4	US-09-618-869-9	Sequence 9, Appli
23	1054	93.4	1068	2	US-08-811-949-52	Sequence 52, Appl
24	1053.8	93.4	1314	2	US-08-811-949-56	Sequence 56, Appl
25	1052.4	93.3	1170	2	US-08-811-949-66	Sequence 66, Appl
26	1050	93.1	1068	6	5223256-3	Patent No. 5223256
27	1046.8	92.8	1065	1	US-08-427-640-5	Sequence 5, Appli
28	1038	92.0	1068	2	US-08-811-949-58	Sequence 58, Appl
29	1035.4	91.8	1068	1	US-08-427-640-3	Sequence 3, Appli
30	1030	91.3	1065	2	US-08-811-949-60	Sequence 60, Appl
31	953.4	84.5	2497	6	5185259-2	Patent No. 5185259
32	924	81.9	1068	2	US-08-427-640-7	Sequence 7, Appli
33	798	70.7	1163	2	US-08-558-269-5	Sequence 5, Appli
34	798	70.7	1163	3	US-09-410-882-5	Sequence 5, Appli
35	625.8	55.5	1727	6	5244676-4	Patent No. 5244676
36	625.6	55.5	1605	6	5244676-1	Patent No. 5244676
37	472	41.8	472	2	US-08-811-949-40	Sequence 40, Appl
38	414	36.7	453	6	5200340-3	Patent No. 5200340
39	267.4	23.7	1724	6	5200340-5	Patent No. 5200340
40	170.8	15.1	1236	1	US-07-957-039A-7	Sequence 7, Appli
41	170.8	15.1	1475	4	US-09-643-597-122	Sequence 122, App
42	170.8	15.1	1475	4	US-09-480-884A-122	Sequence 122, App
43	170.8	15.1	1475	4	US-09-542-615A-122	Sequence 122, App
44	170.8	15.1	1475	4	US-09-606-421B-122	Sequence 122, App
45	170.8	15.1	2301	6	5188829-2	Patent No. 5188829

ALIGNMENTS

RESULT 1
US-08-811-949-62
; Sequence 62, Application US/0811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 62:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1419 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1416
US-08-811-949-62

Query Match   94.4%;   Score 1065;   DB 2;   Length 1419;

Db   1364 TGTACACAAGGTTACCAACTACCTAGACTGGATTCGTGTGACAACATGCGACCG 1416

RESULT 2
US-08-814-412-10
; Sequence 10, Application US/08814412
; Patent No. 6150141
; GENERAL INFORMATION:
;   APPLICANT: Jarrell Ph.D., Kevin A.
;   TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
;   NUMBER OF SEQUENCES: 46
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Choate, Hall & Stewart
;   STREET: 53 State Street
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/814,412
;   FILING DATE: 11-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Jarrell Ph.D., Brenda H.
;   REGISTRATION NUMBER: 39,223
;   REFERENCE/DOCKET NUMBER: 0079571-0040
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617 248 5000
;   TELEFAX: 617 248 4000
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1848 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   IMMEDIATE SOURCE:
;   CLONE: amplified t-PA clone
US-08-814-412-10

Query Match   94.4%; Score 1065; DB 3; Length 1848;
Best Local Similarity 99.5%; Pred. No. 8.6e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
        || ||||||||||||||||||||||||||||||||||||||||||
Db 685 CCCCTGCCTGCTGCTACTTTGGGAATGGGTCAGCCTACCGTG 744

Qy 116 GCACGCACAGCCTCACCGAGTCCGGGTGCCTCCTGCCTCCCGTGCCGTGGAATTCCATGATCCTGA 175
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 745 GCACGCACAGCCTCACCGAGTCCGGGTGCCTCCTGCCTCCCGTGCCGTGGAATTCCATGATCCTGA 804

Qy 176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 805 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 864

Qy 236 ATTACTGCCGGAATCCTGATGGGATGGGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 295
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 865 ATTACTGCCGGAATCCTGATGGGATGGGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 924

Qy 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCACCTGCGGCCTGAGACAGTACA 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 925 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCACCTGCGGCCTGAGACAGTACA 984

Qy 356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 985 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGC 1044

Best Local Similarity 99.5%; Pred. No. 7.7e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
        || ||||||||||||||||||||||||||||||||||||||||||
Db 344 CCCCTGCCTGCTGCTACTTTGGGAATGGGTCAGCCTACCGTG 403

Qy 116 GCACGCACAGCCTCACCGAGTCCGGGTGCCTCCTGCCTCCCGTGCCGTGGAATTCCATGATCCTGA 175
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 404 GCACGCACAGCCTCACCGAGTCCGGGTGCCTCCTGCCTCCCGTGCCGTGGAATTCCATGATCCTGA 463

Qy 176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 464 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 523

Qy 236 ATTACTGCCGGAATCCTGATGGGATGGGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 295
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 524 ATTACTGCCGGAATCCTGATGGGATGGGCCAAGCCTGGTGCCACGTGCTGAAGAACCGCA 583

Qy 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCACCTGCGGCCTGAGACAGTACA 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 584 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCACCTGCGGCCTGAGACAGTACA 643

Qy 356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 644 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 703

Qy 416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTCGTGTGCGGGGCA 475
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 704 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCCGGAGAGCCGGTTCCTCGTGTGCGGGGCA 763

Qy 476 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCACTCGGTCCTTCCGGTGTCCTTGGCGAGGAGAGC 535
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 764 TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCACTCGGTCCTTCCGGTGTCCTTGGCGAGGAGAGC 823

Qy 536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTTGGCGAGGAGAGC 595
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 824 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTTGGCGAGGAGAGC 883

Qy 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 884 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 943

Qy 656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 944 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 1003

Qy 716 TGGTCCGGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1004 TGGTCCGGCACTGTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1063

Qy 776 AGCTCTCCGGCTACGGCAAGCATGAGGCGTTCTCCTTCATTCGGAGGCGCTGAAGG 835
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1064 AGCTCTCCGGCTACGGCAAGCATGAGGCGTTCTCCTTCATTCGGAGGCGCTGAAGG 1123

Qy 836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1124 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1183

Qy 896 CAGTCACCGACAACATGTCTGTGTCTGGGAGGGCTGCGGAGGCGGCGGCCCCAGGCAAACT 955
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1184 CAGTCACCGACAACATGTCTGTGTCTGGGAGGGCTGCGGAGGCGGCGGCCCCAGGCAAACT 1243

Qy 956 TGCACGAGCGCTGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTGTCTGAACGATGGCCCGA 1015
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1244 TGCACGAGCGCTGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTGTCTGAACGATGGCCCGA 1303

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGGTGGGCTGGTGTGAACGAGGATGTCCCGGGTG 1075
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1304 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGGTGGGCTGGTGTGAACGAGGATGTCCCGGGTG 1363

Qy 1076 TGTACACAAGGTTACCAACTACCTAGACTGGATTCGTGTGACAACATGCGACCG 1128
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1364 TGTACACAAGGTTACCAACTACCTAGACTGGATTCGTGTGACAACATGCGACCG 1416

```
OPERATING SYSTEM:  PC-DOS/MS-DOS
SOFTWARE:  PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER:  US/08/811,949
  FILING DATE:  05-MAR-1997
  CLASSIFICATION:  435
ATTORNEY/AGENT INFORMATION:
  NAME:  OBLON, NORMAN F.
  REGISTRATION NUMBER:  24,618
  REFERENCE/DOCKET NUMBER:  18-966-0
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  703-413-3000
  TELEFAX:  703-413-2220
INFORMATION FOR SEQ ID NO:  42:
SEQUENCE CHARACTERISTICS:
  LENGTH:  2101 base pairs
  TYPE:  nucleic acid
  STRANDEDNESS:  double
  TOPOLOGY:  circular
MOLECULE TYPE:  DNA (genomic)
FEATURE:
  NAME/KEY:  CDS
  LOCATION:  25..1710
US-08-811-949-42

Query Match        94.4%;   Score 1065;  DB 2;  Length 2101;
Best Local Similarity  99.5%;   Pred. No. 9.1e-290;
Matches 1068;  Conservative  0;  Mismatches  0;  Indels  5;  Gaps  0;
```

```
Qy     56 CCCAGGCGGCCTTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTG 115
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    638 CCCCTGCCTGCTTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 697

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCTCGCCTCTCCGTGCCTCTCCGTGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    698 GCACGCACAGCCTCACCGAGTCGGGTGCCTCTCGCCTCTCCGTGCCTCTCCGTGAATTCCATGATCCTGA 757

Qy    176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACTGGGCCTGGGCAAAACATA 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    758 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGCACTGGGCCTGGGCACTGGGCCTGGGCAAAACATA 817

Qy    236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    818 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 877

Qy    296 GGCTGACGTGGGAGTACTGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    878 GGCTGACGTGGGAGTACTGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 937

Qy    356 GCCAGCCTCAGTTTCGCATCAAACAGGAGGGCTCTTCGCCGACATGGCGTCTTCGCCGCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    938 GCCAGCCCTCAGTTTCGCATCAAACAGGAGGGCTCTTCGCCGACATGGCGTCCTCGCCGCACCCCTGGC 997

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCGCCGCCCGGAGAGCCGGA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    998 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCGCCGCCCGGAGAGCCGGGCA 1057

Qy    476 TACTCATCAGCTCCTGGATTCTCTGCCGCCCACTGCTTCCGCCGCCCACTGCTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1058 TACTCATCAGCTCCTGGATTCTCTGCCGCCCACTGCTTCCGCCGCCCACTGCTTCCGC 1117

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1118 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGC 1177

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1178 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACA 1237

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCGCTGTGCCCAGGAGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1238 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCCCGCTGTGCCCAGGAGAGCAGCG 1297
```

```
Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCGTCGTGCGGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||
Db   1045 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGTTCGTCGTGCGGGGGGCA 1104

Qy    476 TACTCATCAGCTCCTGGATTCTCTGCCGCCCACTGCTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 TACTCATCAGCTCCTGGATTCTCTGCCGCCCACTGCTTCCGC 1164

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCGAGGAGGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGCGAGGAGGAGC 1224

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACA 1284

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATCGTCGCAGCTGCGAGGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 ATGACATTGCGCTGCTGCAGCTGAAATCGGATCGTCGCAGCTGCGAGGAGCAGCG 1344

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCGGGGACCTGCAGCTGCCGGACTGGACCGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 TGGTCCGCACTGTGTGCCTTCCCCGGGGACCTGCAGCTGCCGGACTGGACCGAGTGTG 1404

Qy    776 AGCTTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG 835
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 AGCTTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGGCTGAAGG 1464

Qy    836 AGGCTCATGTCAGACTGTACCCAGCGCGCTGCACATCACAACATTACTTAACAGAA 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 AGGCTCATGTCAGACTGTACCCAGCGCGCTGCACATCACAACATTACTTAACAGAA 1524

Qy    896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGGCCCCAGGCAAACT 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGGGCCCCAGGCAAACT 1584

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGGAACGATGGCCGCA 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1644

Qy   1016 TGACTTTGGTGGGGCATCATCAGCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1645 TGACTTTGGTGGGGCATCATCAGCTGGGCTGTGGACAGAAGGATGTCCCGGGTG 1704

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACACATGGCGACCG 1128
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1705 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACACATGGCGACCG 1757
```

```
RESULT 3
US-08-811-949-42
; Sequence 42, Application US/08811949
Patent No. 5840533
; GENERAL INFORMATION:
;   APPLICANT:  NIWA, MINEO
;   APPLICANT:  SAITO, YOSHIMASA
;   APPLICANT:  SASAKI, HITOSHI
;   APPLICANT:  HAYASHI, MASAKO
;   APPLICANT:  NOTANI, JOUJI
;   APPLICANT:  KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES:  67
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;     ADDRESS:  P.C.
;     STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;     CITY:  ARLINGTON
;     STATE:  VA
;     COUNTRY:  USA
;     ZIP:  22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
```

FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 335..447
  OTHER INFORMATION: /product= "EGF-like domain"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 448..714
  OTHER INFORMATION: /product= "Kringle-1 domain"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 715..972
  OTHER INFORMATION: /product= "Kringle-2 domain"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 973..2162
  OTHER INFORMATION: /product= "Catalytic domain"

US-08-119-512-3

Query Match   94.4%; Score 1065; DB 1; Length 2162;
Best Local Similarity 99.5%; Pred. No. 9.2e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTG 115
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGAATGGTCAGCCTACCGTG 754

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCGGGTGGCCTCCCGTGGAATTCCATGATCCTGA 175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755 GCACGCACAGCCTCACCGAGTCGGGTGCCTCGGGTGGCCTCCCGTGGAATTCCATGATCCTGA 814

Qy   176 TAGGCAAGTTTACACAGCACAGAACCCCAGTGCCAGCACTGGCCTGGCCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815 TAGGCAAGTTTACACAGCACAGAACCCCAGTGCCAGCACTGGCCTGGCCAAACATA 874

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   875 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 934

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTCCACCTGCCGGCTGCGAGACAGTACA 355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   935 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTCCACCTGCCGGCTGCGAGACAGTACA 994

Qy   356 GCCAGCCTCAGTTTGCGCATCAAAGGAGGGCTCTTCGCCGACAATGCGTCCACCCCTGGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995 GCCAGCCTCAGTTTGCGCATCAAAGGAGGGCTCTTCGCCGACAATGCGTCCACCCCTGGC 1054

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTTCCTGTGTGCGGGGGCA 475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGAGCCGGTTCCTGTGTGCGGGGGCA 1114

Qy   476 TACTCATCAGCTTCCTGCTGGATTCTCTCGCGCGCCACTGCTTCCAGGAGAGGTTTCCGC 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115 TACTCATCAGCTTCCTGCTGGATTCTCTCGCGCGCCACTGCTTCCAGGAGAGGTTTCCGC 1174

Qy   536 CCCACCACTGACGTGATCTTGGCAGAACATACCGGGTGGTCCTCGCGGAGGAGGAGC 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175 CCCACCACTGACGTGATCTTGGCAGAACATACCGGGTGGTCCTCGCGGAGGAGGAGC 1234

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAGCACTTACGACA 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAGCACTTACGACA 1294

Qy   656 ATGACATTGCGCTGCTGAAGCCTGAAATCGGATTCGTCCCGCTGTGCCGAGGAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295 ATGACATTGCGCTGCTGAAGCCTGAAATCGGATTCGTCCCGCTGTGCCGAGGAGCAGCG 1354

Qy   716 TGGTCGGACTGTGTGCCTTCCCCCGGGGAGCTGCAGCTGACCGAGTGTG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355 TGGTCCGACTGTGTGCCTTCCCCCGGGGAGCTGCAGCTGACCGAGTGTG 1414

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCCGGACGCGGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415 AGCTCTCCGGCTACGGCAAGCATGAGGCATGAGGCCTTGTCTCCTTTCTATTCCGGACGCGGCTGAAGG 1474

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGGGCGGGGCCGGACCTGGACCGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1298 TGGTCCGCACTGTGTGCCTTCCCCGGGGCGGGGCCGGACCTGGACCGGAGTGTG 1357

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCATGAGGCCTTGTCTCCTTTCTATTCCGGAGCGGCTGAAGG 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1358 AGCTCTCCGGCTACGGCAAGCATGAGGCATGAGGCCTTGTCTCCTTTCTATTCCGGAGCGGCTGAAGG 1417

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCGCGTCCACATCACAACATTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1418 AGGCTCATGTCAGACTGTACCCATCCAGCGCTGCACATCACAACATTACTTAACAGAA 1477

Qy   896 CAGTCACCCACAACATGCTGTGTGCTCGGAGACACTCGGAGACGCGGGGCGCGGGCCCCAGGCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1478 CAGTCACCGACAACATGCTGTGTGCTCGGAGACACTCGGAGACGCGGGGCGCGGGCCCCAGGCAAACT 1537

Qy   956 TGCACGAGCGCCTGCCAGGGCGATTCGGGAGGCCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1538 TGCACGAGCGCCTGCCAGGGCGATTCGGGAGGCCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1597

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGCGATTCGGGAGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1598 TGACTTTGGTGGGCATCATCAGCTGGGCGATTCGGGAGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1657

Qy   1076 TGTACACAAAGGTTTACCAACTACCTAGACTGGATTCGTGACACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658 TGTACACAAAGGTTTACCAACTACCTAGACTGGATTCGTGACACATGCGACCG 1710

RESULT 4
US-08-119-512-3
; Sequence 3, Application US/08119512
; Patent No. 5498531
GENERAL INFORMATION:
  APPLICANT: Jarrell, Kevin A.
  TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES
  TITLE OF INVENTION: AND REAGENTS
  NUMBER OF SEQUENCES: 3
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: LAHIVE & COCKFIELD
    STREET: 60 State Street
    CITY: Boston
    STATE: MA
    COUNTRY: USA
    ZIP: 02109
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: ASCII (text)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/119,512
    FILING DATE: 10-SEP-1993
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Vincent, Matthew P.
    REGISTRATION NUMBER: 36,709
    REFERENCE/DOCKET NUMBER: HUI-008
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617) 227-7400
    TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2162 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: both
  MOLECULE TYPE: other nucleic acid
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 82..334
    OTHER INFORMATION: /product= "Signal Sequence and
    OTHER INFORMATION: Finger-like domain"

```
;   OTHER INFORMATION: /product= "Kringle-1 domain"
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 715..972
;   OTHER INFORMATION: /product= "Kringle-2 domain"
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 973..2162
;   OTHER INFORMATION: /product= "Catalytic domain"
US-08-488-015B-3

Query Match    94.4%;  Score 1065;  DB 1;  Length 2162;
Best Local Similarity  99.5%;  Pred. No. 9.2e-290;
Matches 1068;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

Qy   56 CCCAGGCGGCCCTGAGCGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
          ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  695 CCCCTGCCTGCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 754

Qy  116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  755 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA 814

Qy  176 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  815 TAGGCAAGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 874

Qy  236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCGTGGTGCCACGTGCTGAAGAACCCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  875 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCGTGGTGCCACGTGCTGAAGAACCCGCA 934

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCCACCTGCCGGCCTGAGACAGTACA 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  935 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCCACCTGCCGGCCTGAGACAGTACA 994

Qy  356 GCCAGCCCTCAGTTTCGCATCAAAGAGAGGGCTCTTCGCCCGACATCGCCTCCCACCCCTGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  995 GCCAGCCCTCAGTTTCGCATCAAAGAGAGGGCTCTTCGCCCGACATCGCCTCCCACCCCTGC 1054

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGAGAGAGCCGGTTCCTGTGCGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1055 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGAGAGAGCCGGTTCCTGTGCGGGGGCA 1114

Qy  476 TACTCATCAGCTCCTGCTTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1115 TACTCATCAGCTCCTGCTTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 1174

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1175 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGGAGC 1234

Qy  596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAAATCGATTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAAATCGATTACGACA 1294

Qy  656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCGGCGGACCTGCAGCTGCCGCAGCGCG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1295 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGGCGGCGGACCTGCAGCTGCCGCAGCGCG 1354

Qy  716 TGGTCCGCCACTGTGTGCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTCCTTTCGATTCGGAGCGGCTGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1355 TGGTCCGCCACTGTGTGCCGCTACGGCAAGCATGAGGCCTTGTCTCCTTCCTTTCGATTCGGAGCGGCTGAGTGTG 1414

Qy  776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCGATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1415 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCGATTCGGAGCGGCTGAAGG 1474

Qy  836 AGGCTCATGTGACCTGTACCCATCCAGCCGCTGCACATCACAACATTACTTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1475 AGGCTCATGTGACCTGTACCCATCCAGCCGCTGCACATCACAACATTACTTTAACAGAA 1534

Qy  896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTGGAGCGCGGCCCCAGCAAACT 955
```

```
Qy  836 AGGCTCATGTGACCTGTACCCATCCAGCCGCTGCACATCACAACATTACTTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1475 AGGCTCATGTGACCTGTACCCATCCAGCCGCTGCACATCACAACATTACTTTAACAGAA 1534

Qy  896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTGGAGCGCGGCCCCAGCAAACT 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1535 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTGGAGCGCGGCCCCAGGCAAACT 1594

Qy  956 TGCACGACGACGCCTGCCAGGCGGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1595 TGCACGACGACGCCTGCCAGGCGGATTCGGGAGGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1654

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCGCCTGGGCGCCTGGGCGCTGTGGACAGAAGGATGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1655 TGACTTTGGTGGGCATCATCAGCTGGGGCGCCTGGGCGCCTGGGCGCTGTGGACAGAAGGATGTCCCGGGTG 1714

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1767
```

```
RESULT 5
US-08-488-015B-3
; Sequence 3, Application US/08488015B
; Patent No. 5780272
; GENERAL INFORMATION:
;   APPLICANT: Jarrell, Kevin A.
;   TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
;   TITLE OF INVENTION: AND REAGENTS
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley, Hoag & Eliot
;     STREET: One Post Office Square
;     CITY: Boston
;     STATE: MA
;     COUNTRY: USA
;     ZIP: 02109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: ASCII (text)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/488,015B
;     FILING DATE: 07-JUN-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Vincent, Matthew P.
;     REGISTRATION NUMBER: 36,709
;     REFERENCE/DOCKET NUMBER: HUV-008.02
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 832-1000
;     TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2162 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: both
;   MOLECULE TYPE: other nucleic acid
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: 82..334
;     OTHER INFORMATION: /product= "Signal Sequence and
; Patent No. 5780272
;     OTHER INFORMATION: Finger-like domain"
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: 335..447
;     OTHER INFORMATION: /product= "EGF-like domain"
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: 448..714
```

```
Db   875 ATTACTGCCGGAATCCTGATGGGATGGCCAAGCCCTGGTGCTGAGAGACCGCA  934

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCTGCTGCTGGCCTGAGACAGTACA  355
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   935 GGCTGACGTGGGAGTACTGTGATGTGCCTGCTGCTGGCCTGAGACAGTACA  994

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCGCACATCGCCCCTGGC  415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTCGCCGCACATCGCCCCCTGC 1054

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGCA  475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1055 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGAGAGCCGGTTCCTGTGCGGGGGGCA 1114

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTGCGGACTCGCTGCTTCCAGGAGAGAGGTTTCGC  535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1115 TACTCATCAGCTCCCGCTGTGGATTCTCTGCGGACTCGCTGCTTCCAGGAGAGAGGTTTCCGC 1174

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAGACATACCGGGTGGTTCCTGCGAGGAGGAGC  595
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1175 CCCACCACCTGACGGTGATCTTGGGCAGAGAACATACCGGGTGTCCCTGCGAGGAGGAGC 1234

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294

Qy   656 ATGACATTGCGCTGCTCAGCTGAAATCGGATTCGTCCCGCGTGTGCCGAGGAGCAGCAGGG  715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1295 ATGACATTGCGCTGCTCAGCTGAAATCGGATTCGTCCCGCGTGTGCCGAGGAGCAGCACGG 1354

Qy   716 TGGTCCGGCACTGTGTGCCTTCCCCCGGCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG  775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1355 TGGTCCGGCACTGTGTGCCTTCGCCCCGGCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1414

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCCTTTCATTCGGAGCGGCTGAGG  835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1415 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCCTTTCATTCGGAGCGGCTGAGG 1474

Qy   836 AGGCTCATGTCAGACTGTACCCATCCGCGTGCACATCACAACATTACTTAACAGAA  895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1475 AGGCTCATGTCAGACTGTACCCATCCGCGTGCACATCACAACATTACTTAACAGAA 1534

Qy   896 CAGTCACCGACAACATGTCGTGTGCTGAGAGACTCGGAGGCGGCGGCCCCAGGCAAACT  955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1535 CAGTCACCGACAACATGTCGTGTGCTGAGAGACTCGGAGGCGGCGGCCCCAGGCAAACT 1594

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1595 TGCACGACGCCTGCCAGGGCGATTCGGAGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1654

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1655 TGACTTTGGTGGGCATCATCAGCTGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGTG 1714

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1715 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1767
```

RESULT 6
US-08-488-015B-25.
; Sequence 25, Application US/08488015B
; Patent No. 5780272
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,015B
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUV-008.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-488-015B-25

Query Match   94.4%; Score 1065; DB 1; Length 2162;
Best Local Similarity   99.5%; Pred. No. 9.2e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy    56 CCCAGGGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGGTG  115
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGGTG  754

Qy   116 GCACGCCACAGCCTCACCCAGTCGGGTGCCTCCTCCGCCTCCTCCGCGGTGGAATTCCATGATCCTGA  175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755 GCACGCCACAGCCTCACCCAGTCGGGTGCCTCCTCCGCCTCCTCCGCGGTGGAATTCCATGATCCTGA  814

Qy   176 TAGGCAAGGTTTACACAGCACAGAGAACCCCAGTCCCAGGCACTGGGCCTGGGCCAACATA  235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815 TAGGCAAGGTTTACACAGCACAGAGAACCCCAGTCCCAGGCACTGGGCCTGGGCCAACATA  874

Qy   236 ATTACTGCCGGAATCCTGATGGGATGGCCAAGCCCTGGTGCTGAGAGAACCGCA  295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 7
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:

```
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;  COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,949
;      FILING DATE:  05-MAR-1997
;      CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  18-966-0
;  TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO: .48:
;  SEQUENCE CHARACTERISTICS:
;      LENGTH:  1314 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;  MOLECULE TYPE:  DNA (genomic)
;  FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1311
US-08-811-949-48
```

```
Query Match          94.3%;  Score 1063.4;  DB 2;  Length 1314;
Best Local Similarity 99.4%;  Pred. No. 2.1e-289;
Matches 1067;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;
```

```
Qy    56  CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239  CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  298

Qy   116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA  175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGGAATTCCATGATCCTGA  358

Qy   176  TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359  TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  418

Qy   236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGTGCTGAAGAACCGCA  295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGTGCTGAAGAACCGCA  478

Qy   296  GGCTGAGCGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGCCTGAGACAGTACA  355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479  GGCTGAGCGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGCGCCTGAGACAGTACA  538

Qy   356  GCCAGCGCACAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCGCACATCGCCTCCCACCCCTGGC  415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539  GCCAGCGCACAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGCCACATCGCCTCCCACCCCTGGC  598

Qy   416  AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCGGAGCGGTTCATGCGTGCGGGGGCA  475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599  AGGCTGCCATCTTTGCCAAGCACAGCAGAGGAGCGGTTCATGCGTGCGGGGGCA  658

Qy   476  TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC  535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659  TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC  718

Qy   536  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC  595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   809  ATTACTGCCGGAATCCTGATGGGGATGCCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC  868
```

```
Db   719  CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC  778

Qy   596  AGAAATTTGAAGTGGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   779  AGAAATTTGAAGTGGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA  838

Qy   656  ATGACATTGCGCTGCTGCAGCTGAAATCGGAATCGGATTCGTCCCGTGCCCGTGCCCAGGAGCAGCG  715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   839  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCGTGCCCAGGAGCAGCG  898

Qy   716  TGGTCCGACTGTGTGCCTTCCCCGGGCGACTGCAGCTGCCGACTGGACGAGTGTG  775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   899  TGGTCCGACTGTGTGCCTTCCCCGGGCGACTGCAGCTGCCGACTGGACGAGTGTG  958

Qy   776  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTTTCCGTGGGCGGCTGAAGG  835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   959  AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTTTCCGTGGGCGGCTGGCTGAAAG  1018

Qy   836  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA  895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1019  AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTAACAGAA  1078

Qy   896  CAGTGCACCGACACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACT  955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1079  CAGTGCACCGACACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACT  1138

Qy   956  TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGCCGCA  1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1139  TGCACGACGCCTGCCAGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGCCGCA  1198

Qy  1016  TGACTTTGGTGGCATCATCAGCTGGACAGTGGATTCGTGACAACATGCGACCG  1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1199  TCACTTTGGTGGGCATCATCAGCTGGACAGTGGATTCGTGACAACATGCGACCG  1258

Qy  1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1259  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1311
```

```
RESULT 8
5200340-1
;Patent No. 5200340
;      APPLICANT:  FOSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
;      TITLE OF INVENTION:  THROMBIN-ACTIVATED TISSUE PLASMINOGEN
;ACTIVATORS
;
;  NUMBER OF SEQUENCES:  34
;  CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/53,412
;      FILING DATE:  22-MAY-1987
;SEQ ID NO:1:
;      LENGTH:  1738
5200340-1
```

```
Query Match          94.3%;  Score 1063.4;  DB 6;  Length 1738;
Best Local Similarity 99.4%;  Pred. No. 2.4e-289;
Matches 1067;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;
```

```
Qy    56  CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  115
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   629  CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG  688

Qy   116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGGGGGCA  175
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   689  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGTGCCTCCCGTGGGGGCA  748

Qy   176  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   749  TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA  808

Qy   236  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGTGCTGAAGAACCGCA  295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   809  ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCCAGTGCTGAAGAACCGCA  868
```

COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1955 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-883-795A-39

Query Match   94.3%;   Score 1063.4;   DB 2;   Length 1955;
Best Local Similarity 99.4%;   Pred. No. 2.5e-289;
Matches 1067; Conservative   0; Mismatches   0; Indels   6; Gaps   0;

```
Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
         ||| || | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  704 CCCCTGCGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 763

Qy  116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCCAGGCACTGGGC 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  764 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCCAGGCACTGGGC 823

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAAACATA 235
         ||||||||||||||||||||||||||||||||||||||||||
Db  824 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCCAAACATA 883

Qy  236 ATTACTGCCGGAATCCTGATCCTGAATGGGATGCCAAGCCTGGTGCCACGTGCCAAGCCGCA 295
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  884 ATTACTGCCGGAATCCTGATCCTGAATGGGATGCCAAGCCTGGTGCCACGTGCCAAGCCGCA 943

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  944 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 1003

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACACATCGCCTCCCACCCCTGGC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1004 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGGCTCTTCGCCGACACATCGCCTCCCACCCCTGGC 1063

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGGCCCGGAGAGCGGTTCCTGTGCGGGGCA 475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1064 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGGCCCGGAGAGCGGTTCCTGTGCGGGGCA 1123

Qy  476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCTCCGCCGCACATACCGGGTGGTCCTGGAGGAGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1124 TACTCATCAGCTCCTGCTGGATTCTCTCTGCTCCGCCGCACATACCGGGTGGTCCTGGAGGAGC 1183

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTCCCTGGCCGAGGAGAGC 595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1184 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTCCCTGGCCGAGGAGAGC 1243

Qy  596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1244 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1303
```

```
Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  869 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 928

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGCACATCGCCTCCCACCCCTGGC 415
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  929 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGCACATCGCCTCCCACCCCTGGC 988

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGGCCCGGAGAGCGGTTCCTGTGCGGGGCA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  989 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGGCCCGGAGAGCGGTTCCTGTGCGGGGCA 1048

Qy  476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCTCCGCCGCACATACCGGGTGGTCCTGGAGGAGC 535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1049 TACTCATCAGCTCCTGCTGGATTCTCTCTGCTCCGCCGCACATACCGGGTGGTCCTGGAGGAGC 1108

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGCAGAACATACCGGGTGGTCCTGGAGGAGC 595
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1109 CCCACCACCTGACGGTGATCTTGGGCAGCAGAACATACCGGGTGGTCCTGGAGGAGC 1168

Qy  596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1169 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1228

Qy  656 ATGACATTGCCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCGCCCAGGCAGCAGCG 715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1229 ATGACATTGCCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCGCCCAGGCAGCAGCG 1288

Qy  716 TGGTTCCGCACTGTGTGCCTTCCCCGGGCGAGCTGCAGCTGGACGGAGTGTG 775
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1289 TGGTTCCGCACTGTGTGCCTTCCCCGGGCGAGCTGCAGCTGGACGGAGTGTG 1348

Qy  776 AGCTCTCCGGCTACGGCCAAGCATCAGCGCCTTGCTCTCTTCTATTCGGAGCGGCTGAAGG 835
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1349 AGCTCTCCGGCTACGGCCAAGCATCAGCGCCTTGCTCTCTTCTATTCGGAGCGGCTGAAGG 1408

Qy  836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAAACAGAA 895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1409 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAAACAGAA 1468

Qy  896 CAGTCACCGACAACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACT 955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1469 CAGTCACCGACAACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCCCAGGCAAACT 1528

Qy  956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTCTGAACGATGGCCCGCA 1015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1529 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGCCCCCTGTGTCTGAACGATGGCCCGCA 1588

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCCTGTGGACAGAAGGAGTGTCCGGGTG 1075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1589 TGACTTTGGTGGGCATCATCAGCTGGGCCTGTGGACAGAAGGAGTGTCCGGGTG 1648

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGAATTCGTGACAACATGCGACCG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1649 TGTACACAAAGGTTACCAACTACCTAGACTGAATTCGTGACAACATGCGACCG 1701
```

RESULT 9
US-08-883-795A-39.
; Sequence 39, Application US/08883795A
; Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcuve, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario

```
Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGGAGCAGCG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGGAGCAGCG 1363

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 TGGTCCGCACTGTGTGCCTTCCCCGGGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1423

Qy    776 AGCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1424 AGCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG 1483

Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1484 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1543

Qy    896 CAGTCACCGACAACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGGCCCCCAGGCCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1544 CAGTCACCGACAACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGGCCCCCAGGCCAAACT 1603

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGTGTGTCTGAACGATGGCCGCA 1663

Qy   1016 TGACTTTGGTGGGGCATCATCAGCTGGGGCGTGGGGCGTGGGGCGTGGGGTCCCGGGGTG 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1664 TGACTTTGGTGGGGCATCATCAGCTGGGGCGTGGGGCGTGGGGCGTGGGGTCCCGGGGTG 1723

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
          |||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1724 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1776
```

RESULT 10
5344773-1
;Patent No. 5344773
;APPLICANT: WEI, CHA-MER;HSIUNG, NANCY;REDDY, VERMURI B.;
;LEMONTT, JEFFREY F.;DACKOWSKI, WILLIAM;DOUGLAS, RICHARD;
;COLE, EDWARD S.;PURCELL JR., RICHARD D.;LAU, DAVID TAI-YUI
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
;ACTIVATOR PRODUCED BY RECOMBIANT DNA
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/782,686
;   FILING DATE: 01-OCT-1985
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 656,770
;   FILING DATE: 01-OCT-1984
;SEQ ID NO:1:
; LENGTH: 2457
5344773-1

Query Match   94.3%;   Score 1063.4;   DB 6;   Length 2457;
Best Local Similarity   99.4%;   Pred. No. 2.7e-289;
Matches 1067;   Conservative   0;   Mismatches   6;   Indels   0;   Gaps   0;

```
Qy     56 CCCAGGCGGCCCTCTGAGGGAAACAGTGACTGCTACTTTGGGGATGGGTCAGCCTACCGTG 115
          ||| ||    |||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    627 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 686

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTGCCTCCGTGTCCCGTGGAATTCCATGATCCTGA 175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 GCACGCACAGCCTCACCGAGTCGGGTGCCTGCCTCCGTGTCCCGTGGAATTCCATGATCCTGA 746

Qy    176 TAGGCAAGGTTTACACAGCACAGCACAGAACCCAGTCGCCAGGCACTGGGCCTGGGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 TAGGCAAGGTTTACACAGCACAGCACAGAACCCAGTCGCCAGGCACTGGGCCTGGGCAAACATA 806

Qy    236 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCAAGCCCTGCTGTCCACGTGCTGTGAAGAACCGCA 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807 ATTACTGCCGGAATCCTGATGGGGATGGGGATGCCAAGCCCTGCTGTCCACGTGCTGTGAAGAACCGCA 866

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCACCTGCGCGCCTGAGACAGTACA 355
```

```
Db    867 GGCTGACGTGGGAGGTACTGTGATGTGCGATGTGCTGCTTCCTGCTCCACTGCGACAGTACA 926
          ||||||||||||| |||||||||||||
Qy    356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTTCCACCCCTGGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTTCCACCCCTGGC 986

Qy    416 AGGCTGCCATCTTGCCAAGCACAGACAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    987 AGGCTGCCATCTTGCCAAGCACAGACAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 1046

Qy    476 TACTCATCAGCTCCTGCTGATTCTCTGCTGCCCACTGCTTCCGGAGGAGGTTTCCGC 535
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1047 TACTTCATCAGCTCCTGCTGATTCTCTGCTGCCCACTGCTTCCGGAGGAGGTTTCCGC 1106

Qy    536 CCCACCACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGAGGAGGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1107 CCCACCACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGGAGGAGGAGC 1166

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1167 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1226

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGGAGCAGCG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1227 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGGAGCAGCG 1286

Qy    716 TGGTCCGCACTGTGTGCCTTCCCCGGGGCGACTGCCGAGCTGCCGGACTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1287 TGGTCCGCACTGTGTGCCTTCCCCGGGGCGACTGCCGAGCTGCCGGACTGGACGGAGTGTG 1346

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1347 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTCTATTCGGAGCGGCTGAAGG 1406

Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1466

Qy    896 CAGTCACCGACAACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 CAGTCACCGACAACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGGGCCCCAGGCAAACT 1526

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGTGTGTCTGAACGATGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCTGTGTGTCTGAACGATGCCGCA 1586

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCATCAGCTGGGGCATCATCAGCTGGGGCATCAGCTCCGGGGTG 1075
          |||||||||||| |||||||||||||||||||
Db   1587 TGACTTTGGTGGGCATCATCAGCTGGGGCATCAGCTGGGGCATCATCAGCTGGGGCATCAGCTCCGGGGTG 1646

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
          |||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1647 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1699
```

RESULT 11
US-08-286-740-1
; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
;   APPLICANT: Crowley, Craig W.
;   TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
;   TITLE OF INVENTION: HOST CELLS
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:

```
MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
COMPUTER:  IBM PC compatible
OPERATING SYSTEM:  PC-DOS/MS-DOS
SOFTWARE:  patin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  US/08/286,740
  FILING DATE:  05-AUG-1994
  CLASSIFICATION:  435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
ATTORNEY/AGENT INFORMATION:
  NAME:  Lee, Wendy M.
  REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:  798
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  415/225-1994
  TELEFAX:  415/952-9881
  TELEX:  910/371-7168
INFORMATION FOR SEQ ID NO:  1:
SEQUENCE CHARACTERISTICS:
  LENGTH:  7360 bases
  TYPE:  nucleic acid
  STRANDEDNESS:  double
  TOPOLOGY:  linear
US-08-286-740-1
```

```
Query Match          94.3%;  Score 1063.4;  DB 1;  Length 7360;
Best Local Similarity 99.4%;  Pred. No. 4.3e-289;
Matches 1067;  Conservative  0;  Mismatches  6;  Indels  0;  Gaps  0;
```

```
Qy     56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2235 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 2294

Qy    116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTCCTCCGTGGAATTCCATGATCCTGA 175
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2295 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCGTCCTCCGTGGAATTCCATGATCCTGA 2354

Qy    176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2355 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 2414

Qy    236 ATTACTGCCGGAATCCTGATGGGGATGGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2415 ATTACTGCCGGAATCCTGATGGGGATGGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 2474

Qy    296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCACCTGCGGCCTGAGCAGTACA 355
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2475 GGCTGACGTGGGAGTACTGTGATGTGCCCTCACCTGCGGCCTGAGACAGTACA 2534

Qy    356 GCCAGCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGC 415
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2535 GCCAGCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGC 2594

Qy    416 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGACAGAGGCGGTTCCTGTCGCGGGGCA 475
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2595 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGACAGAGGCGGTTCCTGTGTCGCGGGGCA 2654

Qy    476 TACTCATCAGCTGCCTGCTGGATTCTCTCGCCCACTGCTTCTGCGCCCACTGCTTCTGCGC 535
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2655 TACTCATCAGCTGCCTGCTGGATTCTCTCGCCCACTGCTTCTGCGCCCACTGCTTCTGCGC 2714

Qy    536 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGGC 595
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2715 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 2774

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2775 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 2834

Qy    656 ATGACAATTGCGCTGCTGCCAGCTGAAATCGGATTCGTCCCGCGTGCTGCCCAGGAGCAGCCG 715
             |||||||||||||||||||||||||||||||||||||||||||
Db   2835 ATGACAATTGCGCTGCTGCCAGCTGAAATCGGATTCGTCCCGCGTGCTGCCCAGGAGCAGCCG 2894

Qy    716 TGGTCCCGCACTGTGTGCCTTCCCCCGGCGGACTGCCGCAGCTGCCGCCGACTGACGGAGTGTG 775
             |||||||||||||||||||||||||||||||||||||||||||
Db   2895 TGGTCCCGCACTGTGTGCCTTCCCCCGGCGGACTGCCGCAGCTGCCGCCGACTGACGGAGTGTG 2954

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCGGAGCGCGGCTGAAGG 835
             |||||||||||||||||||||||||||||||||||||||||||
Db   2955 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCGGAGCGGCGGCTGAAGG 3014

Qy    836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
             |||||||||||||||||||||||||||||||||||||||||||
Db   3015 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 3074

Qy    896 CAGTCACCGACAACATGTCGTGTGCTGGACACTCGGAGGCGCGGGCCCCCCAGGCCAAACT 955
             |||||||||||||||||||||||||||||||||||||||||||
Db   3075 CAGTCACCGACAACATGTCGTGTGCTGGACACTCGGAGGCGCGGGCCCCCAGGCCAAACT 3134

Qy    956 TGCACGACGCCTGCCAGGGGGCGATTCGGGAGGCCGATTCGGGAGGCCCCCCTGGTGTCTGAACGATGGCCGCA 1015
             |||||||||||||||||||||||||||||||||||||||||||
Db   3135 TGCACGACGCCTGCCAGGGGGCGATTCGGGAGGCCGATTCGGGAGGCCCCTGTGTCTGAACGATGGCCGCA 3194

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGTGGACAGAAGGATGTCCCGGGTG 1075
             |||||||||||||||||||||||||||||||||||||||||||
Db   3195 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGTGGACAGAAGGATGTCCCGGGTG 3254

Qy   1076 TGTACAAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
             |||||||||||||||||||||||||||||||||||||||||||
Db   3255 TGTACAAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 3307
```

```
RESULT 12
PCT-US95-09576-1
; Sequence 1, Application PC/TUS9509576
GENERAL INFORMATION:
;   APPLICANT:  GENENTECH, INC.
;   TITLE OF INVENTION:  METHOD FOR SELECTING HIGH-EXPRESSING
;   TITLE OF INVENTION:  HOST CELLS
;   NUMBER OF SEQUENCES:  4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Genentech, Inc.
;     STREET:  460 Point San Bruno Blvd
;     CITY:  South San Francisco
;     STATE:  California
;     COUNTRY:  USA
;     ZIP:  94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  patin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  PCT/US95/09576
;     FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/286740
;     FILING DATE:  05-AUG-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Lee, Wendy M.
;     REGISTRATION NUMBER:  00,000
;     REFERENCE/DOCKET NUMBER:  798PCT
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  415/225-1994
;     TELEFAX:  415/952-9881
;     TELEX:  910/371-7168
; INFORMATION FOR SEQ ID NO:  1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  7360 bases
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
PCT-US95-09576-1
```

Qy   1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3255 TGTACACCAAGGTTACCAACTACCTAGACTACTTAGACTGGATTCGTGACAACATGCGACCG 3307

RESULT 13
US-08-811-949-44
; Sequence 44, Application US/08811949
; Patent No. 5840533
GENERAL INFORMATION:
;   APPLICANT: NIWA, MINEO
;   APPLICANT: SAITO, YOSHIMASA
;   APPLICANT: SASAKI, HITOSHI
;   APPLICANT: HAYASHI, MASAKO
;   APPLICANT: NOTANI, JOUJI
;   APPLICANT: KOBAYASHI, MASAKAZU
;   TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/811,949
;   FILING DATE: 05-MAR-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 18-966-0
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 44:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1065
US-08-811-949-44

Query Match   94.1%; Score 1062; DB 2; Length 1068;
Best Local Similarity 100.0%; Pred. No. 4.8e-289;
Matches 1062; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   67 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4 TCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACGCACAGC 63

Qy   127 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 CTCACCGAGTCGGGTGCGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 123

Qy   187 TACACAGCACAGAACCCCAGTGCCACCACTGGGCCTCGGCAAACATAATTACTGCCGG 246
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 TACACAGCACAGAACCCCAGTGCCACCACTGGGCCTCGGCAAACATAATTACTGCCCG 183

Qy   247 AATCCTGATGGGGATGCCAAGCCCTGGTGTCCACGTGCTGCTGAAGAACCGCAGGCTGACGTGG 306
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 AATCCTGATGGGGATGCCAAGCCCTGGTGTCCACGTGCTGAAGAACCGCAGGCTGACGTGG 243

Query Match   94.3%; Score 1063.4; DB 5; Length 7360;
Best Local Similarity 99.4%; Pred. No. 4.3e-289;
Matches 1067; Conservative   0; Mismatches   6; Indels   0; Gaps   0;

Qy   56 CCCAGGCGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
          |||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2235 CCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 2294

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2295 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 2354

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGGCCTCGGCAAACATA 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2355 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGGCCTCGGCAAACATA 2414

Qy   236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGTCCACGTGCTGAAGAACCGCA 295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2415 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGTCCACGTGCTGAAGAACCGCCA 2474

Qy   296 GGCTGACGTGGGAGTACTGTGATGTGCCCCTGCTCCACCTGCGCGCCTGAGACAGTACA 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2475 GGCTGACGTGGGAGTACTGTGATGTGCCCCTGCTCCACCTGCGCGCCTGAGACAGTACA 2534

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2535 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGC 2594

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCGCCCGGAGAGCCGTTCTGTGCGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2595 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCGCCCGGAGAGCCGTTCTGTGCGGGGCA 2654

Qy   476 TACTCATCAGCTCCTGCTGGATTCTCTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2655 TACTCATCAGCTCCTGCTGGATTCTCTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGC 2714

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGAGC 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2715 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGAGC 2774

Qy   596 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2775 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 2834

Qy   656 ATGACATTGCGCGTGCTGCAGCTGAAATCGGAATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2835 ATGACATTGCGCGTGCTGCAGCTGAAATCGGAATTCGTCCCGCTGTGCCCAGGAGAGCAGCG 2894

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGAACTGCCAGCTGCCGGACTGGACGGAGTGTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2895 TGGTCCGCACTGTGTGCCTTCCCCCGGCGGAACTGCCAGCTGCCGGACTGGACGGAGTGTG 2954

Qy   776 AGCTCTCCGGCTACGGCAAGCATGGAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2955 AGCTCTCCGGCTACGGCAAGCATGGAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGCCGCTGAAGG 3014

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3015 AGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 3074

Qy   896 CAGTGACCGACAACATGCTGTGTGCTGAGACACTCCGAGCCGGCGGCCCAGGCAAACT 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3075 CAGTGACCGACAACATGCTGTGTGCTGAGACACTCCGAGCCGGCGGCCCAGGCAAACT 3134

Qy   956 TGCAGCGACGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGTGTCTGAACGATGGCCGCA 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3135 TGCAGCGACGCCTGCCAGGGCGATTCGGGAGGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 3194

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGGGCCTGGGGCTGTGTGACACAGAAGGATGTCCCGGGTG 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3195 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGGGCCTGGGGCTGTGTGACACAGAAGGATGTCCCGGGTG 3254

```
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/811,949
     FILING DATE: 05-MAR-1997
     CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
     NAME: OBLON, NORMAN F.
     REGISTRATION NUMBER: 24,618
     REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
     TELEPHONE: 703-413-3000
     TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 50:
     SEQUENCE CHARACTERISTICS:
     LENGTH: 1314 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: double
     TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
FEATURE:
     NAME/KEY: CDS
     LOCATION: 1..1311
US-08-811-949-50
```

```
Query Match           94.1%;  Score 1061.8;  DB 2;  Length 1314;
Best Local Similarity 99.3%;  Pred. No. 6e-289;
Matches 1066;  Conservative 0;  Mismatches 7;  Indels 0;  Gaps 0;
```

```
Qy   56 CCCAGGCCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTCAGCCTACCGTG 115
        |||    ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  239 CCCCTGCCTGCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTCAGCCTACCGTG 298

Qy  116 GCACGCAGAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCCGTGA 175
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 GCACGCAGAGCCTCACCGAGTCGGGTGCCTCCGTGCCTCCGTGCCTCCGTGA 358

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGGCCTGGGCCAAACATA 235
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCACTGGGCCTGGGCCAAACATA 418

Qy  236 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCTGTGCCACGTGCTGAAGAACCGCA 295
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  419 ATTACTGCCGGAATCCTGATGGGATGGGATGCCAAGCCTGTGCCACGTGCTGAAGAACCGCA 478

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTCTCCACCTGCGGCCTGAGACAGTACA 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  479 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCGTCTCCACCTGCGGCCTGAGACAGTACA 538

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCGACATCGCCCACCCTGGC 415
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  539 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCGACATCGCCCACCCTGGC 598

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  599 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 658

Qy  476 TACTCATCAGCTCCTGCTGGATTCTCTGCTGCTTCCAGGAGAGGTTTCCGC 535
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  659 TACTCATCAGCTCCTGCTGGATTCTCTGCTGCTTCCAGGAGAGGTTTCCGC 718

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  719 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 778
```

```
Qy  307 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 366
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244 GAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAG 303

Qy  367 TTTCGGCATCAAAGGAGGGGCTCTTCGCCGACATGCGCCAGGCTGCCATC 426
        |||||||||||||||||||||||||||||| |||||||||||||||||||
Db  304 TTTCGGCATCAAAGGAGGGGCTCTTCGCCCACATGCGCCAGGCTGCCATC 363

Qy  427 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGC 486
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364 TTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGC 423

Qy  487 TCCTGCTGGATTCTTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 546
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424 TCCTGCTGGATTCTTCTGCCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCACCACCTG 483

Qy  547 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 606
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAA 543

Qy  607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 666
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 603

Qy  667 CTGCTGCAGCTGAAATCGGATTCGTCCCGGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 726
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGCCCAGGAGAGCAGCGTGGTCCGCACT 663

Qy  727 GTGTGCCTTCCCCCGGCGCAGCTGCAGCTGCCGACTGCAGCTGAGCGCGGCTGAAGGAGGCTCATGTC 786
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664 GTGTGCCTTCCCCCGGCGCAGCTGCAGCTGCCGACTGCAGCTGAGCGCGGCTGAAGGAGGCTCATGTC 723

Qy  787 TACGGCAAGCATGAGGCCCTTGCTCCTTTCTATTCGGAGCGCGCTGAAGGAGGCTCATGTC 846
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  724 TACGGCAAGCATGAGGCCCTTGCTCCTTTCTATTCGGAGCGCGCTGAAGGAGGCTCATGTC 783

Qy  847 AGACTGTACCCATCCAGCCGCTGCACATCACACACATTTACTTAACAGAACAGTCACCGAC 906
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784 AGACTGTACCCATCCAGCCGCTGCACATCACACACATTTACTTAACAGAACAGTCACCGAC 843

Qy  907 AACATGCGTGTGTCTGGAGAGCACTCGGAGCGCGGCGGGCCCCCGTGTGTCTGAACGATGGCCGATGACTTTGGTG 966
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  844 AACATGCGTGTGTCTGGAGAGCACTCGGAGCGCGGCGGGCCCCCGTGTGTCTGAACGATGGCCGATGACTTTGGTG 903

Qy  967 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGTGTGTCTGAACGATGGCCGATGACTTTGGTG 1026
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  904 TGCCAGGGGCGATTCGGGAGGAGGCCCCCTGTGTGTCTGAACGATGGCCGATGACTTTGGTG 963

Qy 1027 GGCATCATCAGCTGGGGCGTGGGCCTGGGCGTGTGTGGACAGAAGAGATGTCCCGGGGTGTGTACACAAAG 1086
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  964 GGCATCATCAGCTGGGGCGTGGGCCTGGGCGTGTGTGGACAGAAGAGATGTCCCGGGGTGTGTACACAAAG 1023

Qy 1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
        ||||||||||||||||||||||||||||||||||||||||||
Db 1024 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1065
```

```
RESULT 14
US-08-811-949-50
; Sequence 50, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
;  APPLICANT: NIWA, MINEO
;  APPLICANT: SAITO, YOSHIMASA
;  APPLICANT: SASAKI, HITOSHI
;  APPLICANT: HAYASHI, MASAKO
;  APPLICANT: NOTANI, JOUJI
;  APPLICANT: KOBAYASHI, MASAKAZU
;  TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
;  NUMBER OF SEQUENCES: 67
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
```

```
                              TELEPHONE:   703-413-3000
                              TELEFAX:     703-413-2220
                           INFORMATION FOR SEQ ID NO:    38:
                           SEQUENCE CHARACTERISTICS:
                              LENGTH:  1974 base pairs
                              TYPE: nucleic acid
                              STRANDEDNESS: double
                              TOPOLOGY: circular
                           MOLECULE TYPE: DNA (genomic)
                           FEATURE:
                              NAME/KEY:  CDS
                              LOCATION:  3..1583
US-08-811-949-38

Query Match           94.1%;    Score 1061.8;  DB 2;  Length 1974;
Best Local Similarity 99.3%;    Pred. No. 7e-289;
Matches 1066;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;
```

```
Qy    56 CCCAGGCGGCCCTCGAGCGAAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
             |||   || ||||||||||||||||||||||||||||||||||||||||||||||||
Db   511 CCCCTGCCTCGTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 570

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGTGGAGCACTGGGCCTGGAATTCCATGATCCTGA 175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   571 GCACGCACAGCCTCACCGAGTCGGGTGCCTCCCGTGGAGCACTGGGCCTGGAATTCCATGATCCTGA 630

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGACACTGGGCCTGGCCTGGCAAACATA 235
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   631 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCAGGACACTGGGCCTGGCAAACATA 690

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCAGGCCCTGGTGCCACGTGCCAGGCCGCA 295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   691 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCCAGGCCGCA 750

Qy   296 GGCTGACGTGGGAGTACTGTGATGCGCCCTCTCGTGCTCAACGGCCTGAGACAGTACA 355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   751 GGCTGACGTGGGAGTACTGTGATGCGCCCTCTCGTGCTCAACGGCCTGAGACAGTACA 810

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGGCCTCTTCGCCGACATCGCTCCCACCCCTGGC 415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   811 GCCAGCCTCAGTTTCGCATCATAGGAGGGCCTCTTCGCCGACATCGCTCCCACCCCTGGC 870

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   871 AGGCTGCCATCTTTGCCAAGCACAGGAGAGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 930

Qy   476 TACTCATCAGCTCCTGCTTCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   931 TACTCATCAGCTCCTGCTTCTGGATTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGC 990

Qy   536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   991 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1050

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1051 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1110

Qy   656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCGCCAGGAGAGCAGCG 715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1111 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTCGCCAGGAGAGCAGCG 1170

Qy   716 TGGTCCGCACTGTGTGCCTTCCCCGGACTGCAGCTGCCGGAGTGTG 775
             ||||||||||||||||||||||||||||||||||||||||||||||
Db  1171 TGGTCCGCACTGTGTGCCTTCCCCGGACTGCAGCTGCCGGAGTGTG 1230

Qy   776 AGCTCTCCGGCTACGGCCAAGCATGAGGCCTTGTTCTATTCGGAGCGGCTGAAGG 835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1231 AGCTCTCCGGCTACGGCCAAGCATGAGGCCTTGTTCTATTCGGAGCGGCTGAAGG 1290

Qy   836 AGGCTCATGTCAGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAA 895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1291 AGGCTCATGTCAGACTGTACCCATCCAGCGCGCTGCACATCACAACATTTACTTAACAGAA 1350
```

Additional alignment blocks:

```
Qy   896 CAGTCACCGACAACATGCTGTGTCTGAGAGACACTCGGAGCGGCGGACCTGCAGGCAAACT 955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1079 CAGTCACCGACAACATGCTGTGTCTGAGAGACACTCGGAGCGGCGGACCTGCAGGCAAACT 1138

Qy   956 TGCACGAGCGGCCTGCCAGGGCGGATTCGGGAGGGAGGCCCCTGGTGTGTCTGAACGATGGCCCGCA 1015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1139 TGCACGAGCGGCCTGCCAGGGCGGATTCGGGAGGGAGGCCCCTGGTGTGTCTGAACGATGGCCCGCA 1198

Qy  1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGCGTTGTGGACAGAAGGATGTCCCGGGTG 1075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1199 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGCGTTGTGGACAGAAGGATGTCCCGGGTG 1258

Qy  1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1259 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG 1311
```

```
RESULT 15
US-08-811-949-38
;  Sequence 38, Application US/08811949
;  Patent No. 5840533
GENERAL INFORMATION:
   APPLICANT: NIWA, MINEO
   APPLICANT: SAITO, YOSHIMASA
   APPLICANT: SASAKI, HITOSHI
   APPLICANT: HAYASHI, MASAKO
   APPLICANT: NOTANI, JOUJI
   APPLICANT: KOBAYASHI, MASAKAZU
   TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
   NUMBER OF SEQUENCES:  67
   CORRESPONDENCE ADDRESS:
      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
      ADDRESSEE:  P.C.
      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
      CITY:  ARLINGTON
      STATE:  VA
      COUNTRY:  USA
      ZIP:  22202
   COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/811,949
      FILING DATE:  05-MAR-1997
      CLASSIFICATION:  435
   ATTORNEY/AGENT INFORMATION:
      NAME:  OBLON, NORMAN F.
      REGISTRATION NUMBER:  24,618
      REFERENCE/DOCKET NUMBER:  18-966-0
   TELECOMMUNICATION INFORMATION:
```

```
Qy    896  CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGCGGGCCCCAGGCAAACT   955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1351  CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGCGCGGGCCCCAGGCAAACT  1410

Qy    956  TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA  1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411  TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA  1470

Qy   1016  TGACTTTGTGGGCATCATCAGCTGGGGCCTCGGGCTGTTGGACAGAAGGATGTCCCGGGTG  1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471  TGACTTTGTGGGCATCATCAGCTGGGGCCTGGGGCCTGTGGACAGAAGGATGTCCCGGGTG  1530

Qy   1076  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1531  TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG  1583
```

Search completed: August 19, 2003, 08:35:06
Job time : 90.2804 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 318.307 Seconds
(without alignments)
7930.701 Million cell updates/sec

Title: US-09-987-455-2
Perfect score: 1128
Sequence: 1 atgaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 150479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1128	100.0	1128	11	US-09-987-455-2	Sequence 2, Appli
2	1128	100.0	1128	11	US-09-987-455-5	Sequence 5, Appli
3	1063.4	94.3	1689	9	US-09-969-271-6	Sequence 6, Appli
4	1063.4	94.3	2509	14	US-10-193-656-7	Sequence 7, Appli
5	1063.4	94.3	2519	9	US-09-969-271-5	Sequence 5, Appli
6	1062	94.1	1065	11	US-09-987-455-4	Sequence 4, Appli
7	1062	94.1	1065	11	US-09-987-455-7	Sequence 7, Appli
8	1061.8	94.1	2641	10	US-09-974-298-144	Sequence 144, App
9	227.2	20.1	329	12	US-10-007-926A-433	Sequence 433, App
10	191.6	17.0	1212	10	US-09-880-503-15	Sequence 15, Appl
11	188	16.7	1415	14	US-10-198-846-12748	Sequence 12748, A
12	170.8	15.1	1475	9	US-09-735-705-122	Sequence 122, App
13	170.8	15.1	1475	10	US-09-850-716A-122	Sequence 122, App
14	170.8	15.1	1475	10	US-09-897-778-122	Sequence 122, App
15	170.8	15.1	1475	11	US-09-466-396A-122	Sequence 122, App
16	170.8	15.1	1475	12	US-10-117-982-122	Sequence 122, App
17	170.8	15.1	1475	12	US-10-101-510-159	Sequence 159, App
18	170.8	15.1	2336	12	US-09-971-392-46	Sequence 46, Appl
19	170.8	15.1	2336	12	US-10-101-510-458	Sequence 458, App
20	167.6	14.9	1296	13	US-10-076-421-1	Sequence 1, Appli
21	167.6	14.9	2294	9	US-09-735-705-123	Sequence 123, App
22	167.6	14.9	2294	10	US-09-850-716A-123	Sequence 123, App
23	167.6	14.9	2294	10	US-09-897-778-123	Sequence 123, App
24	167.6	14.9	2294	11	US-09-466-396A-123	Sequence 123, App
25	167.6	14.9	2294	12	US-10-117-982-123	Sequence 123, App
26	167.6	14.9	2344	12	US-10-101-510-383	Sequence 383, App
27	166	14.7	1236	10	US-09-880-503-12	Sequence 12, Appl
28	166	14.7	2304	12	US-10-301-822-160	Sequence 160, App
29	166	14.7	2304	14	US-10-171-311-183	Sequence 183, App
30	163	14.5	474	9	US-09-864-761-1682	Sequence 1682, Ap
31	162	14.4	2486	14	US-10-106-698-1989	Sequence 1989, Ap
32	159.8	14.2	831	10	US-09-880-503-14	Sequence 14, Appl
33	159.8	14.2	972	10	US-09-880-503-16	Sequence 16, Appl
34	139	12.3	461	12	US-09-918-995-10405	Sequence 10405, A
35	123	10.9	1959	12	US-10-101-510-743	Sequence 743, App
36	123	10.9	2035	12	US-10-101-510-448	Sequence 448, App
37	123	10.9	2048	11	US-09-858-909-1	Sequence 1, Appli
38	120.4	10.7	2036	11	US-09-954-456-552	Sequence 552, App
39	120.4	10.7	2036	10	US-09-880-107-1612	Sequence 1612, Ap
40	111.4	9.9	116	9	US-09-864-761-18438	Sequence 18438, A
41	105	9.3	457	11	US-09-918-995-11188	Sequence 11188, A
c 42	89.4	7.9	372	10	US-09-920-300A-237	Sequence 237, App
c 43	89.4	7.9	372	13	US-10-03-528-237	Sequence 237, App
44	88	7.8	442	11	US-09-918-995-12715	Sequence 12715, A
45	83.8	7.4	1230	9	US-09-879-792-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-987-455-2
; Sequence 2, Application US/0987455
; Publication No. US2003004972A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2:
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match   100.0%; Score 1128; DB 11; Length 1128;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGTTTCGCTACCGTGGCCCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGTTTCGCTACCGTGGCCCAG 60

Qy   61 GCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGGCACG 120

US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match   100.0%;   Score 1128;   DB 11;   Length 1128;
Best Local Similarity   100.0%;
Matches 1128;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGTGTTCGTACCGTGGCCAG   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGTGTTCGTACCGTGGCCAG   60

Qy   61 GCGGCCTCACCGAGTGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTGGCACG  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCGGCCTCACCGAGTGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTGGCACG  120

Qy  121 CACAGCCTCACCGAGTGCCTCCGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGATAGGC  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CACAGCCTCACCGAGTGCCTCCGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGATAGGC  180

Qy  181 AAGTTTACACAGCACAGAACCCCAGTGCCAAGCACTGGCCTGGGCCTGGGCCAAACATAATTAC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAGTTTACACAGCACAGAACCCCAGTGCCAAGCACTGGCCTGGGCCTGGGCCAAACATAATTAC  240

Qy  241 TGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGAAGAACCCAGGCTG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGAAGAACCCAGGCTG  300

Qy  301 ACGTGGGAGTACTGTGATGTGCCCTCTGCGGCCTGAGACAGTACAGCCAG  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ACGTGGGAGTACTGTGATGTGCCCTCTGCGGCCTGAGACAGTACAGCCAG  360

Qy  361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGCAGGCT  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGCAGGCT  420

Qy  421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTC  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTC  480

Qy  481 ATCAGCTCCGTGCTGGATTCTCTGCGCGCTCCTGCGCGCTTCCAGGAGAGGTTCCGCCCCAC  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ATCAGCTCCGTGCTGGATTCTCTGCGCGCTCCTGCGCGCTTCCAGGAGAGGTTCCGCCCCAC  540

Qy  541 CACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAA  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAA  600
```

RESULT 2

```
Db   61 GCGGCCTCACCGAGTGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCGTGGCACG  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 CACAGCCTCACCGAGTGCCTCCGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGATAGGC  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CACAGCCTCACCGAGTGCCTCCGTGCCTCCGTGCCTCCGTGGAATTCCATGATCCTGATAGGC  180

Qy  181 AAGTTTACACAGCACAGAACCCCAGTGCCAAGCACTGGCCTGGGCCTGGGCCAAACATAATTAC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAGTTTACACAGCACAGAACCCCAGTGCCAAGCACTGGCCTGGGCCTGGGCCAAACATAATTAC  240

Qy  241 TGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGAAGAACCCAGGCTG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGAAGAACCCAGGCTG  300

Qy  301 ACGTGGGAGTACTGTGATGTGCCCTCTGCGGCCTGAGACAGTACAGCCAG  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ACGTGGGAGTACTGTGATGTGCCCTCTGCGGCCTGAGACAGTACAGCCAG  360

Qy  361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGCAGGCT  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCCTGGCAGGCT  420

Qy  421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTC  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATACTC  480

Qy  481 ATCAGCTCCGTGCTGGATTCTCTGCGCGCTCCTGCGCGCTTCCAGGAGAGGTTCCGCCCCAC  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ATCAGCTCCGTGCTGGATTCTCTGCGCGCTCCTGCGCGCTTCCAGGAGAGGTTCCGCCCCAC  540

Qy  541 CACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAA  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CACCTGACGGTGATCTTGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGCAGAAA  600

Db  601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC  660

Db  661 ATTGCGGCTGTGCAGTCGAAATCGGATTCGTCCCGTGCCCGGAGCAGCAGCGTGTC  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 ATTGCGGCTGTGCAGTCGAAATCGGATTCGTCCCGTGCCCGGAGCAGCAGCGTGTC  720

Db  721 CGACTGTGTGCCTTCCCCCGGGGACCTGCAGCTGCGCCGGACCTGGACGTGGGACTGGGAGCTC  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721 CGACTGTGTGCCTTCCCCCGGGGACCTGCAGCTGCGCCGGACCTGGACGTGGGACTGGGAGCTC  780

Db  781 TCCGGCTACGGCAAGCATGAGGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGAGGAGGCT  840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781 TCCGGCTACGGCAAGCATGAGGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGAGGAGGCT  840

Db  841 CATGTCAGACTGTACCCATCCAGCGCTGCACATCAGACAATCACAACAGTC  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841 CATGTCAGACTGTACCCATCCAGCGCTGCACATCAGACAATCACAACAGTC  900

Db  901 ACCGACAACACATGCTGTGTGCTCGGAGCACTCGGAGCGCGGCCCCAGGCAAACTTGCAC  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901 ACCGACAACACATGCTGTGTGCTCGGAGCACTCGGAGCGCGGCCCCAGGCAAACTTGCAC  960

Db  961 GACGCCTGCCAGGCGATTCGGGAGGGCCGGTGTGTCTGGAACGATGGCCGCATGACT 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961 GACGCCTGCCAGGCGATTCGGGAGGGCCGGTGTGTCTGGAACGATGGCCGCATGACT 1020

Db 1021 TTGGTGGGCATCATCAGCTGGGGCGCCTGGGCTGTGGACAGAGGATGTCCCGGGGTGTAC 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 TTGGTGGGCATCATCAGCTGGGGCGCCTGGGCTGTGGACAGAGGATGTCCCGGGGTGTAC 1080

Db 1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
        |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

```
Qy  601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC 660

Qy  661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGAGCAGCGTGTC 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGAGCAGCGTGTC 720

Qy  721 CGACTGTGTGCCTTCCCCGGCGGAACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CGACTGTGTGCCTTCCCCGGCGGAACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTC 780

Qy  781 TCCGGCTACGGCAAGCATGAGCGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCT 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 TCCGGCTACGGCAAGCATGAGCGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCT 840

Qy  841 CATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACACAATTACTTAACAGAACAGTC 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CATGTCAGACTGTACCCATCCAGCCGCGCTGCACATCACACAATTACTTAACAGAACAGTC 900

Qy  901 ACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCCGCGGCCCCCAGGCAAACTTGCAC 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 ACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCCGCGGCCCCCAGGCAAACTTGCAC 960

Qy  961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGAACGATGGCCGCCATGACT 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 GACGCCTGCCAGGGCGATTCGGGAGGCCCCCCTGGTGTGTCTGAACGATGGCCGCCATGACT 1020

Qy 1021 TTGGTGGGCATCATCAGCTGGGCTGGGCCTGGGCGTGTGGACAGAAGGAGGATGTGTCCCGGGTGTGTAC 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TTGGTGGGCATCATCAGCTGGGCTGGGCCTGGGCGTGTGGACAGAAGGAGGATGTGTCCCGGGTGTGTAC 1080

Qy 1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
```

RESULT 3
US-09-969-271-6
; Sequence 6, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6

Query Match   94.3%; Score 1063.4; DB 9; Length 1689;
Best Local Similarity   99.4%; Pred. No. 0; Gaps 0;
Matches 1067; Conservative 0; Mismatches 6; Indels 0;

```
Qy   56 CCCAGGCCGGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 115
        |||||  ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  614 CCCCTGCCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 673

Qy  116 GCACGCCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  674 GCACGCCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 733

Qy  176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGCCACTGGGCCTGGGCAAACATA 235
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  734 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGCCACTGGGCCTGGGCAAACATA 793

Qy  236 ATTACTGCCGGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
        ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  794 ATTACTGCCGAATCCTGATGGGGATGGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 853

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCAACTGGGCCTGCAGACAGTACA 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  854 GGCTGACGTGGGAGTACTGTGATGTGCCCTGCTCCAACTGGGCCTGCAGACAGTACA 913

Qy  356 GCCAGGCCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTCCCACCCCTGC 415
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  914 GCCAGGCCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTCCCACCCCTGC 973

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGCAGGAGTCGCCCGAGAGGCGGTTCTGTGCGGGGGCA 475
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  974 AGGCTGCCATCTTTGCCAAGCACAGCAGGAGTCGCCCGAGAGGCGGTTCTGTGCGGGGGCA 1033

Qy  476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCAGGAGAACAGTC 535
        ||||||||||||||||||||||||||||||||||||||||||||
Db 1034 TACTCATCAGCTCCTGCTGGATTCTCTCTGCCAGGAGAACAGTC 1093

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCCCTGGCGGAGGAGC 595
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1094 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCCCTGGCGGAGGAGC 1153

Qy  596 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1154 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1213

Qy  656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCG 715
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1214 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGCAGCG 1273

Qy  716 TGGTCCGCACTGTGTGCCTTCCCCGGCGGACTGCAGCTGCCGGACTGGACGGAGTGTG 775
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1274 TGGTCCGCACTGTGTGCCTTCCCCGGCGGACTGCAGCTGCCGGACTGGACGGAGTGTG 1333

Qy  776 AGCTCTCCGGCTACGGCAAGCAAGCATGAGCGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 835
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1334 AGCTCTCCGGCTACGGCAAGCAAGCATGAGCGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGG 1393

Qy  836 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACAATTACTTAACAGAA 895
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1394 AGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACACAATTACTTAACAGAA 1453

Qy  896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCCCCCAGGCAAACT 955
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1454 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCCCCCAGGCAAACT 1513

Qy  956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCCA 1015
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1514 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCCA 1573

Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCATGGGCCTGGGCGTGTGGACAGAAGGAGATGTCCCGGGTG 1075
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1574 TGACTTTGGTGGGCATCATCAGCTGGGCATGGGCCTGGGCGTGTGGACAGAAGGAGATGTCCCGGGTG 1633

Qy 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
        |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 1634 TGTACACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1686
```

RESULT 4
US-10-193-656-7
; Sequence 7, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656

```
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: US 60/304,461
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/304,490
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/305,182
;  PRIOR FILING DATE: 2001-07-13
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.1
;  SEQ ID NO 7
;  LENGTH: 2509
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;  DATABASE ACCESSION NUMBER: GenBank / X07393
;  DATABASE ENTRY DATE: 1995-03-27
;  RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match          94.3%;  Score 1063.4;  DB 14;  Length 2509;
Best Local Similarity 99.4%;  Pred. No. 0;
Matches 1067; Conservative  0; Mismatches  6; Indels  0; Gaps  0;

Qy      56 CCCAGGCGGCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
           |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     667 CCCCTGCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 726

Qy     116 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCGTCCCGTGGAATTCCATGATCCTGA 175
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     727 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCGTCCCGTGGAATTCCATGATCCTGA 786

Qy     176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     787 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 846

Qy     236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGTCGCCACGTGCTGAGAACCGCA 295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     847 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGTCGCCACGTGCTGAGAACCGCA 906

Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     907 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 966

Qy     356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGCTCTTCGGCCGACATCGCCTCCCACCCCTGGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     967 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGCTCTTCGGCCGACATCGCCTCCCACCCCTGGC 1026

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGAGGACGGTTCGCTGGCGGAGAGAGCCAGCCA 475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1027 AGGCTGCCATCTTTGCCCAAGCACAGAGGACGGTTCGCTGGCGGAGAGAGCCAGCCA 1086

Qy     476 TACTCATCAGCTCGCTGGAATTCTCTCTGCGCCCACTGCTTCCGTGGCTTCCCGC 535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1087 TACTCATCAGCTCGCTGGAATTCTCTCTGCGCCCACTGCTTCCGTGGCTTCCCGC 1146

Qy     536 CCCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGTCCCTGGCGAGGAGAGC 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1147 CCCACCACCTGACGGTGATCTTGGGCAGAAACATACCGGGTGTCCCTGGCGAGGAGAGC 1206

Qy     596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACA 655
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1207 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGATGACACTTACGACA 1266

Qy     656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCGGAGAGCAGCG 715
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1267 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCGGAGAGCAGCG 1326

Qy     716 TGGTTCCGCACTGTGTGCCTTCCCCCGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1327 TGGTTCCGCACTGTGTGCCTTCCCCCGGGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1386

Qy     776 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTCTTTCTATTCGGAGCGGCTGAAGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1387 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTCTTTCTATTCGGAGCGGCTGAAGG 1446

Qy     836 AGGCTCATGTGCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1447 AGGCTCATGTGCAGACTGTACCATCCAGCCGCTGCACATCACAACATTACTTACTTAACAGAA 1506

Qy     896 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGACGCGGGGCCCCAGGCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1507 CAGTCACCGACAACATGCTGTGTCTGGAGACACTCGGAGACGCGGGGCCCCAGGCAAACT 1566

Qy     956 TGCACGACGCCTGCCAGGGCGATTCGGCAGGCGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1567 TGCACGACGCCTGCCAGGGCGATTCGGCAGGCGGCCCCCTGTGTGTCTGAACGATGGCCGCA 1626

Qy    1016 TGACTTTGGTGGGGCATCATCAGCTGGGCCTGGGCTGGCTGTGGGACAGAAGGATGTCCCGGGTG 1075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1627 TGACTTTGGTGGGGCATCATCAGCTGGGCCTGGGCTGGCTGTGGGACAGAAGGATGTCCCGGGTG 1686

Qy    1076 TGTACACAAAGGTTACCAACTACTAGACTGGATTCGTGACAACATGCGACCG 1128
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1687 TGTACACCAAGGTTACCAACTACTAGACTGGATTCGTGACAACATGCGACCG 1739

RESULT 5
US-09-969-271-5
;  Sequence 5, Application US/09969271
;  Patent No. US20020098179A1
;  GENERAL INFORMATION:
;  APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
;  APPLICANT: Pfizer Limited (GB and EP (GB) only)
;  TITLE OF INVENTION: Pharmaceutical Combinations
;  FILE REFERENCE: PCS10951APME
;  CURRENT APPLICATION NUMBER: US/09/969,271
;  CURRENT FILING DATE: 2001-10-01
;  PRIOR APPLICATION NUMBER: GB 0025473.0
;  PRIOR FILING DATE: 2000-10-17
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 5
;  LENGTH: 2519
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-969-271-5

Query Match          94.3%;  Score 1063.4;  DB 9;  Length 2519;
Best Local Similarity 99.4%;  Pred. No. 0;
Matches 1067; Conservative  0; Mismatches  6; Indels  0; Gaps  0;

Qy      56 CCCAGGCGGCCTCGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 115
           |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     690 CCCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGTCAGCCTACCGTG 749

Qy     116 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCGTCCCGTGGAATTCCATGATCCTGA 175
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     750 GCACGCACAGCCTCACCGAGTCGGTGCCTCCTGCGTCCCGTGGAATTCCATGATCCTGA 809

Qy     176 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     810 TAGGCAAGGTTTACACAGCACAGAACCCAGTGCCAGGCACTGGGCCTGGGCAAACATA 869

Qy     236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGTCGCCACGTGCTGAAGAACCGCA 295
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     870 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCTGTCGCCACGTGCTGAAGAACCGCA 929

Qy     296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     930 GGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 989

Qy     356 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGCTCTTCGGCCGACATCGCCTCCCCTGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     990 GCCAGCCTCAGTTTCGCATCAAAGGAGGAGGCTCTTCGGCCGACATCGCCTCCCCTGGC 1049

Qy     416 AGGCTGCCATCTTTGCCAAGCACAGAGGAGGTCGCCCGGAGAGCCGTTCCTGTGCGGGGCA 475
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: coding
;  OTHER INFORMATION: sequence for K2S protein
US-09-987-455-4

Query Match        94.1%;   Score 1062;  DB 11;  Length 1065;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1062;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     67 TCTGAGGGAAACAGTGACTGCTACTTTCGGGAATGGTCAGCCTACCGTGGCACGCACAGC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTGAGGGAAACAGTGACTGCTACTTTCGGGAATGGTCAGCCTACCGTGGCACGCACAGC 60

Qy    127 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCACCGAGTCGGGTGCCTCCTGCCTCCCGTCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT 120

Qy    187 TACACAGCACAGAACCCCAGTGCCAGTGACTGGCCTGGCAAACATAATTACTGCCGG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACAGCACAGAACCCCAGTGCCAGTGACTGGCCTGGCAAACATAATTACTGCCGG 180

Qy    247 AATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCGGAGGCTGACGTGG 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATCCTGATGGGATGCCAAGCCCTGTGCCACGTGCTGAAGAACCGGAGGCTGACGTGG 240

Qy    307 GAGTACTGTGATGTGCCTCTTGCCACTGAGACAGTAGCAGTACAGCCAGCCTCAG 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGTACTGTGATGTGCCTCTTGCCACTGAGACAGTAGCAGTACAGCCAGCCTCAG 300

Qy    367 TTTCGCATCAAAGGAGGGCTCTTGCCGCACATCGCCTCCACCCTGGCAGGCTGCCATC 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCGCATCAAAGGAGGGCTCTTGCCGCACATCGCCTCCACCCTGGCAGGCTGCCATC 360

Qy    427 TTTGCCAAGCACAGGAGTCGCCGGACTCGGAGCGCGGTTCGTGTGCGGGGGCATACTCATCAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTGCCAAGCACAGGAGTCGCCGGACTCGGAGCGCGGTTCGTGTGCGGGGGCATACTCATCAGC 420

Qy    487 TCCTGCTCTGATTCTCTCGCGCCCACTGCTTCCGCCCACCACCTG 546
          ||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCTGCTCTGATTCTCTCGCGCCCACTGCTTCCGCCCACCACCTG 480

Qy    547 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAGAAATTTGAA 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGAGAAATTTGAA 540

Qy    607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600

Qy    667 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGGTGGTCCGGCACT 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCGGTGGTCCGGCACT 660

Qy    727 GTGTGCCTTCCCCGCGGCAGACTGGACTGCAGCTGCGCGGACTGACGAGTGTGAGCTCTCCGGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTGTGCCTTCCCCGCGGCAGACTGGACTGCAGCTGCGCGGACTGACGAGTGTGAGCTCTCCGGC 720

Qy    787 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGCTGAACGAGGAGGCTCATGTC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGCTGAACGAGGAGGCTCATGTC 780

Qy    847 AGACTGTACCCATCCAGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGACTGTACCCATCCAGCCGCTGACATCACAACATTTACTTAACAGAACAGTCACCGAC 840

Qy    907 AACATGCTGTGTCGTGAGACACTCGAGGCGGCGGCCCCAGGCAAACTTGCACGACGCC 966
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACATGCTGTGTCGTGAGACACTCGAGGCGGCGGCCCCAGGCAAACTTGCACGACGCC 900

Qy    967 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGCCGATGACTTTGGTG 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TGCCAGGGCGATTCGGGAGGCCCCCTGGTGTCTGAACGATGCCGATGACTTTGGTG 960

Db   1050 AGGCTGCCATCTTTGCCAAGCACAGGAGTCGCCGGACTCGGAGCGCGGTTCCTGTGCGGGGGCA 1109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    476 TACTCATCAGCTCCTGCTCTGATTCTCTCGCGCCCACTGCTTCCGCCCACTGCTTCCGC 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 TACTCATCAGCTCCTGCTCTGATTCTCTCGCGCCCACTGCTTCCGCCCACTGCTTCCGC 1169

Qy    536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 CCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGAGGAGC 1229

Qy    596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1289

Qy    656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCG 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGCAGCG 1349

Qy    716 TGGTCCGGCACTGTGTGCCTTCCCCGCGGCAGCTGGACTGCAGCTGGACGGAGTGTG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 TGGTCCGGCACTGTGTGCCTTCCCCGCGGCAGCTGGACTGCAGCTGGACGGAGTGTG 1409

Qy    776 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGCTGAAGG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410 AGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGCTGAAGG 1469

Qy    836 AGGCTCATGTCAGACTGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 AGGCTCATGTCAGACTGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1529

Qy    896 CAGTCACCGACAACATGCTGTGTCGTGAGAGACACTCGGAGGCGGCGGCCCCAGGCAAACT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 CAGTCACCGACAACAACATGCTGTGTCGTGAGAGACACTCGGAGGCGGCGGCCCCAGGCAAACT 1589

Qy    956 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGATTCGGGAGGCGCTGGGGTCTGTGTGTCTGAACGATGGCCGCA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590 TGCACGACGCCTGCCAGGGCGATTCGGGAGGCGATTCGGGAGGCGCTGGGGTCTGTGTGTCTGAACGATGGCCGCA 1649

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGGTCTGTGTGGCTGTGGCTGTGACAGAAGGATGTCCCGGGTG 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGGTCTGTGTGGCTGTGGCTGTGACAGAAGGATGTCCCGGGTG 1709

Qy   1076 TGTACACAAAGGTTACCACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1710 TGTACACCAAGGTTACCACTACCTAGACTGATTCGTGACAACATGCGACCG 1762

RESULT 6
US-09-987-455-4
;  Sequence 4, Application US/09987455
;  Publication No. US2003004972A1
;  GENERAL INFORMATION:
;  APPLICANT: Aranya Manosroi
;  APPLICANT: Jiradej Manosroi
;  APPLICANT: Chatchai Tayapiwatana
;  APPLICANT: Friedrich Goetz
;  APPLICANT: Rolf-Guenther Werner
;  TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;  TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;  FILE REFERENCE: 0652.2190001
;  CURRENT APPLICATION NUMBER: US/09/987,455
;  CURRENT FILING DATE: 2001-11-14
;  PRIOR APPLICATION NUMBER: 60/268,574
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: GB 0027779.8
;  PRIOR FILING DATE: 2000-11-14
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
;
;  SEQ ID NO 4
;  LENGTH: 1065
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
```

Qy   1027 GGCATCATCAGCTCGGGCGGGGCTGTGGGCGTCGGGGCTGTGGGCGTGTCCGGGTGTGTACACAAAG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCATCATCAGCTCGGGCGGGGCTGTGGGCGTCGGGGCTGTGGGCGTGTCCGGGTGTGTACACAAAG 1020

Qy   1087 GTTACCAACTACCTAGACTGGATTCGTGACAGCAACATGCGACCG 1128
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTACCAACTACCTAGACTGGATTCGTGACAGCAACATGCGACCG 1062

RESULT 7
US-09-987-455-7
; Sequence 7, Application US/0987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
US-09-987-455-7

Query Match   94.1%; Score 1062; DB 11; Length 1065;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1062; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   67 TCTGAGGGAAACAGTGACTGACTACTTTGGAATGGGTCAGCCTACCGTGGCACGCACAGC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCTGAGGGAAACAGTGACTGACTACTTTGGAATGGGTCAGCCTACCGTGGCACGCACAGC 60

Qy   127 CTCACCGGTCGGGTGCCCTTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCACCGGTCGGGTGCCCTTCCTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT 120

Qy   187 TACCAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 246
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TACCAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGCAAACATAATTACTGCCGG 180

Qy   247 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGTGACGTGG 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCCGCAGGTGACGTGG 240

Qy   307 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGCCTCAGAGACAGCAGTACAGCCTCAG 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAGTACTGTGATGTGCCCCTCCTGCTCCACCTGCGGCCTCAGAGACAGCAGTACAGCCTCAG 300

Qy   367 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360

Qy   427 TTTGCCAAGCACAGGAGAGTGCCCGAGAGCCGTTCCTGTGCGGGGGCATACTCATCAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTTGCCAAGCACAGGAGAGTGCCCGAGAGCCGTTCCTGTGCGGGGGCATACTCATCAGC 420

Qy   487 TCCTGCTGGATTCTCTCGCCGCCCACCGTTTCCGCCCCACCACCTG 546
          |||||||||||||||||||||||||||||||||||||||

Db   421 TCCTGCTGGATTCTCTCGCCGCCCACTGCGTTCCGCCCCACCACCTG 480

Qy   547 ACGGTGATCTTGGGCAGAACATACCGGTGGTCCTGGTCCTGCGGAGGAGGAGCAGAAATTTGAA 606
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGTGATCTTGGGCAGAACATACCGGTGGTCCTGGTCCTGCGGAGGAGGAGCAGAAATTTGAA 540

Qy   607 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATGACATTGCG 666
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACATGACATTGCG 600

Qy   667 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGGACGCAGCGTGGTCCGCACT 726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGGACGCAGCGTGGTCCGCACT 660

Qy   727 GTGTGCCTTCCCCGGCGGACCTGCCAGCTGCCGGACTGGCGAGTGTGAGCTCTCCGGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GTGTGCCTTCCCCGGCGGACCTGCCAGCTGCCGGACTGGCGAGTGTGAGCTCTCCGGC 720

Qy   787 TACGGCAAGCATGAGGCCTTGTCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TACGGCAAGCATGAGGCCTTGTCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTC 780

Qy   847 AGACTGTACCCATCCAGCGCGTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 906
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGACTGTACCCATCCAGCGCGTGCACATCACAACATTTACTTAACAGAACAGTCACCGAC 840

Qy   907 AACATGCTGTGTGCTGGAGACACTGGAGCGGGGCCCCAAACTTGCACGACGCC 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACATGCTGTGTGCTGGAGACACTGGAGCGGGGCCCCAAACTTGCACGACGCC 900

Qy   967 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGCGCGATGACTTTGGTG 1026
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TGCCAGGGCGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGCGCGATGACTTTGGTG 960

Qy   1027 GGCATCATCAGCTCGGGCTGGGCGTCGGGGCTGTGGGCGTGTGTACACAAAG 1086
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCATCATCAGCTCGGGCTGGGCGTCGGGGCTGTGGGCGTGTGTACACAAAG 1020

Qy   1087 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1128
          ||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTACCAACTACCTAGACTGGATTCGTGACAACATGCGACCG 1062

RESULT 8
US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US2002015263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002015263A1 1001470CB1
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

Query Match   94.1%; Score 1061.8; DB 10; Length 2641;
Best Local Similarity   99.3%; Pred. No. 0;
Matches 1066; Conservative   0; Mismatches   7; Indels   7; Gaps   0;

RESULT 9
US-10-007-926A-433
; Sequence 433, Application US/10007926A
; Publication No. US20030143539A1
GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 433
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (82)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (159)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (264)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (278)
; OTHER INFORMATION: a, t, c or g
US-10-007-926A-433

Query Match   20.1%; Score 227.2; DB 12; Length 329;
Best Local Similarity   97.4%; Pred. No. 5.9e-62;
Matches 229; Conservative   0; Mismatches   6; Indels   0; Gaps   0;

Qy   894 AACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGGCCCCAGGCAAA 953
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   1 AACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGGCGGGCCCCAGGNAAA 60

Qy   954 CTTGCACGACGCCTGCCAGGGCGATTCGGCGGACATTCGGGAGGCGCCCCTGGTGTGTCTGAACGATGGCCG 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTTGCACGACGCCTGCCAGGGNATTCGGCGGACATTCGGGAGGCGCCCCTGGTGTGTCTGAACGATGGCCG 120

Qy   1014 CATGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTCTGGACAGAGGATGTCCCGG 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CATGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTCTGGACAGAAGGATGTCCCGG 180

Qy   1074 TGTGTACACAAAGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGTGTACACCAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 235

RESULT 10
US-09-880-503-15
; Sequence 15, Application US/09980503
; Patent No. US20020131964A1

Qy   56 CCCAGGCGGCCTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCTACCGTG 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 CCCCTGCCTGCTCTGAGGGAAAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTG 869

Qy   116 GCACGCACAGCCTCACCGAGTCGGGTGCGGGTGCCTCCTGCCTCCTGCCTCCGTGAATTCCATGATCCTGA 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 GCACGCACAGCCTCACCGAGTCGGGTGCGGGTGCCTCCTGCCTCCTGCCTCCGTGAATTCCATGATCCTGA 929

Qy   176 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 TAGGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 989

Qy   236 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCA 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 ATTACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGGCA 1049

Qy   296 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 GGCTGACGTGGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACA 1109

Qy   356 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCACATCGCCTCCCACCCCTGC 415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 GCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCACATCGCCTCCCACCCCTGGC 1169

Qy   416 AGGCTGCCATCTTTGCCAAGCACAGGAGCGCCGTTCGTGTGCGGGGGCA 475
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 AGGCTGCCATCTTTGCCAAGCACAGGAGCGCCGTTCCTGTGCGGGGGGCA 1229

Qy   476 TACTCCATCAGCTCCTGCTGGATTCTCTCGCTGCCTTCCAGGAGAGAGTTTCCGC 535
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 TACTCCATCAGCTCCTGCTGGATTCTCTCTGCCGCCCACTGCCTTCCAGGAGAGAGTTTCCGC 1289

Qy   536 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGTGGTCCTGGCGAGGAGC 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 CCCACCACCTGACCGGTGATCTTGGGCAGAACATACCGGTGGTCCTGGCGAGGAGC 1349

Qy   596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1409

Qy   656 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410 ATGACATTGCGCTGCTCGAGCTGAAATCGGATTCGTCCCGCTGCCCAGGAGAGCAGCG 1469

Qy   716 TGGTCCGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 775
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 TGGTCCGACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGGACGGAGTGTG 1529

Qy   776 AGCTCTCCGGCTACGGCAAGCATGAAGCATGAGCGGCTTGTCTCCTTCTATTCGGAGGCGCTGAAGG 835
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 AGCTCTCCGGCTACGGCAAGCATGAAGCATGAGCGGCCTTGTCTCCTTCTATTCGGAGGCGGCTGAAGG 1589

Qy   836 AGGCTCATGTCAGACTGTACCGATCCAGCCGCTGCACATCACAACATTACTTAACAGAA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590 AGGCTCATGTCAGACTGTACCGATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1649

Qy   896 CAGTGCACGACAACATGCTGTGTCTGGAGACGGCGGCCCCCTCGGAGGCGGCCCAAACT 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 CAGTCACCGACAACATGCTGTGTGCTGGAGACGGCGGCCCCCTCGGAGGCGGCCCAGCCAAACT 1709

Qy   956 TGCACGACGCCTGCCAGGGCGATTCGGCGAGGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1710 TGCACGACGCCTGCCAGGGCGATTCGGCGAGGGCCCCTGGTGTGTCTGAACGATGGCCGCA 1769

Qy   1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTCTGGACAGAAGGATGTCCCGGGTG 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1770 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTCTGGACAGAAGGATGTCCCGGGTG 1829

Qy   1076 TGTACACAAAGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1128
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1830 TGTTACACCAAGGTTACCAACTACCTAGACTGATTCGTGACAACATGCGACCG 1882

; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 1212
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-880-503-15

Query Match   17.0%;   Score 191.6;   DB 10;   Length 1212;
Best Local Similarity   52.3%;   Pred. No. 2.1e-50;
Matches   549; Conservative   0; Mismatches 459; Indels   42; Gaps   4;

Qy   84 CTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACCAGCCTCACCGAGTCGGGTGC 143
Db   147 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAGGCCAGCACTGACACCATGGGCGCG 206

Qy   144 CTCCTGCCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACGCACAGAGACCC 203
Db   207 GCCCTGCCTGCCCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCATGCCCACAGATC 266

Qy   204 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263
Db   267 TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCGGAAACCCAGACAACCGGAG 326

Qy   264 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCTGACGTGGGAGTACTGTGATGTGCC 323
Db   327 GCGACCCTGTGTGTATGTGCAGTGGGCCTTAAACGCCGTTGTCCAAAGACTCTGAGTGGTCA 386

Qy   324 CTCCTGCTCCACCTGGCGCGGCCTGAGACGTACAGCCA----------GCCTCAGTT 368
Db   387 TGACTGCGCAGATGGAAAATTAAAATTTCAGTGTGGCCAAAAGACTCTGAGGCCCCGCTT 446

Qy   369 TCGGCATCAAAGAGAGGGGCTCTTCGCCCGACATCGCCTCCACCCTGCAGCCTGCCATCTT 428
Db   447 TAAGATTATTGGGGAGAATTCACCACCATCGAACCAGCCCGGTTGCGGTTGCGGCCATCTA 506

Qy   429 TGCCAAGCACAGGAGGTCGCCCGGAGACGCGGTTCCTGTGCGGGGCATACTCATCAGCTC 488
Db   507 CAGGAGGCACCGGGGGGCTCGTCA---CCTACGTGTGTGGAGGCAGCCTCATCAGCCC 563

Qy   489 CTGCTGGATTCTCTCTGCGCCCACTGCGCTTCCAGGAGAGGTTTCCGCCCCACCTGAC 548
Db   564 TTGCTGGGTGATCAGCGCCACACACTGCTTCATTGATTACCCAAAGAAGAGGACTACACAT 623

Qy   549 GGTGATCTTGGGCAGAACATACCCGGTGCCCGGAGGAGGCAGCAGAAATTTGAAGT 608
Db   624 CGTCTACACCGGGTCGCTCAAGGCTTACTCCAACACGCAAGGGGAGATGAAGTTTGAGGT 683

Qy   609 CGAAAAATACATTGTCATAAGGAATTCGATAGATGACA------CTTACGACAATGACAT 662
Db   684 GGAAAAACCTCATCCTACACAAAGGACTACAGCGCTTGCGACACGCTGACACCACAACGACAT 743

Qy   663 TGCGGCTGCTGCAGCTGAAATCGGATTCGTCTCCCGCTGTGCCCAGGAGCAGCGCTGGGTCCG 722
Db   744 TGCCTTGCCTGAAGATCCGTTCCAAGGAGGCAGGTGTCGCAGCCATCCGGACTACTACA 803

Qy   723 CACTGTGTGCCTTCCCCCGGCGACCTGCAGCTGCCGGGGACTGAGCGGAGTGTGTGAGCTCTC 782
Db   804 GACCCATCTGCCTGCCCTCGATGTATAACAGCTGTCCCAGTTTGGCACAACGCTGTGAGATCAC 863

Qy   783 CGGCTACGGCAAGCATGAGGGCCTTGCTGTCTTCTTTTCTATTCGGAGCGGCTGAAGGAGGGCTCA 842
Db   864 TGGCTTTTGGAAAAGAGAGAATTCTACCGACTATCTCTATCCGGACGTATAATGACTGT 923

RESULT 11
US-10-198-846-12748
; Sequence 12748, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12748
;   LENGTH: 1415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1_2, 3, 1414, 1415
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12748

Query Match   16.7%;   Score 188;   DB 14;   Length 1415;
Best Local Similarity   91.1%;   Pred. No. 3.1e-49;
Matches   255; Conservative   0; Mismatches   20; Indels   5; Gaps   5;

Qy   56 CCCAGGCGGCCTCTGAGGGAAACAGTGACTGCTACTTTTGGGAATGGGCTACTTGGGATCACCGTG 115
Db   895 CCCCTGCCTGCTCGTCTGAGGGAAAACAGTGACTG-TACTTTGGGAATGGGCTACTTGGGATCACCGTG 953

Qy   116 GCACGCAGCAGCCTCACCGAGTCGGGTGCCGGTCCCT-GCCTCCCGTGGAATTCCATGATCCTG 174
Db   954 GCACGCACAG-CTCACCGAGTCGGGTGCCGGTCCCTGCCTCCCGTGGAATTCCATGATCCTG 1012

Qy   175 ATAGGCAAGGTTTACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACAT 234
Db 1013 ATAGGCAAGGTTTACACAGCACAGAGACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACAT 1072

Qy   235 AATTACTGCCGGAATCCTGATGGGCCCACGTGCCAAGCCCTGGTGCCACGTGCTGAA-GAACCG 293
Db 1073 AAATTACTGCCGGAAT-CTGATGGGCCCACGTGCCAAGCCCTGGTGCCACGTGCTGAACGAACCG 1131

Qy   294 CAGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCC 333

```
Db   1132 CAGGCTGACGTGGGAGTACTGTGATGTGGCCTCTGTCACC 1171

RESULT 12
US-09-735-705-122
; Sequence 122, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Fan, Liqun
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT: Hosken, Nancy
;  APPLICANT: Fanger, Gary R.
;  APPLICANT: Li, Samuel X.
;  APPLICANT: Wang, Aijun
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Henderson, Robert A.
;  APPLICANT: McNeill, Patricia D.
;  APPLICANT: Fanger, Neil
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C14
;  CURRENT APPLICATION NUMBER: US/09//735,705
;  CURRENT FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 419
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;  LENGTH: 1475
;  TYPE: DNA
;  ORGANISM: Homo sapien
US-09-735-705-122

Query Match             15.1%;  Score 170.8;  DB 9;  Length 1475;
Best Local Similarity   51.3%;  Pred. No. 9.6e-44;
Matches  551;  Conservative    0;  Mismatches  457;  Indels  66;  Gaps  4;

Qy    84 CTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGCCTCACCGAGTCGGGTGC 143
Db   287 CTGCTATGAGGGAATGGTCACTTTTACCGAGGAAAAGGCCAGCACTGACACCATGGGCCG 346

Qy   144 CTCCTGCCTCCCGTGGAATTCCATGATCCTGCACTGTCCTTCCTTCAGCAAACATAATTACTGC 203
Db   347 GCCCTGCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCATGCCCACAGATC 406

Qy   204 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263
Db   407 TGATGCTCTTCAGCTGGGGCCTGGGGAAACATAATTACTGCCGAGAACCCCAGAACCCGGAG 466

Qy   264 CAAGCCCTGGTGCCACGTGCTGAAGAACCCAGCTGGAGTACTGTGATGTGCC 323
Db   467 GCGACCCTGGTGCTATGTGCAGTGGCCTAAAGCCGCTGTCCAAAGAGTGCATGGTGCA 526

Qy   324 CTCCTGCTCCACCTGCGGCCTGAG- 347
Db   527 TGACTGCCAGATGGAAAAAAGCCCTCTCCTCCTCCAGAAGAGAATTAAAATTTCAGTGTGG 586

Qy   348 --ACAGTACAGCCAGCCAGTTCGCATCAAAGGAGGGCTCTTCGCCCGACATCGCCTC 404
Db   587 CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAATTCACCACCATGGGCCAGGAA 646

Qy   405 CCAGCCCTGGCCAGGCTGCCATCTTGCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCCT 464
Db   647 CCAGCCCTGGTTTGCGGCCATCCACAGAGGACCGGGG---GGGCTCTGTCACCTACGT 703

Qy   465 GTGCGGGGGCATACTCATCAGCGTCCTGCTCCTGTCCCACTGCTTCCTCAGGA 524
Db   704 GTGTGGAGGCAGCCTCATCAGCCCCTGCTGGGTGATCAGCGCCACACTGCTTCATTGA 763

Qy   525 GAGGTTTCCGCCCACCACCACCTGACGCGTGATCTTGGGCAGAACATACCCGGGTGTCCCTGG 584
Db   764 TTACCCAAGAGGAGGACTACATCGTCTCACCTGGTCGCTGCTCAAGGCTTAACTCCAACAC 823
```

```
Qy   585 CGAGGAGGAGCAGCAAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGA 644
Db   824 GCAAGGGGGAGATGAAGTTTGAGGTGGAAAACCTCATCCTACACAGGACTACAGCGCTGA 883

Qy   645 CA------CTTACGACAATGACATTGGCTGCTGCTGCTGAACTCGGATTCGTCCCGCTG 698
Db   884 CACGCTTGCTCACCACACGACATTGCCTTGCTGTGAAGATCCGTTCCAAGGAGGAGGCAGGTG 943

Qy   699 TGCCCAGGAGCAGCGTGTCCGCACTGTGTCGTTGCTTCCCCCGGGGACCTGCAGCTGCC 758
Db   944 TGGCCAGCCGCCATCCGGACTATACAGCCATCGCTGCCGCTGCGATGTATAACGATCCCCA 1003

Qy   759 GGACTGGACGGAGTGTGAGCTCTCCGGTACGGCAAGCATGAGGCCTTGTCTCCTTTCTA 818
Db  1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAATTCTACCGACTATCTCTA 1063

Qy   819 TTCCGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACA 878
Db  1064 TCCGGAGCAGCTGAAGATGACTGTGTTGTGAAGCTGGATTCCCACCGGAGTGTCAGCAGCC 1123

Qy   879 ACATTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCCGGAGCGG 938
Db  1124 CCACTACTACGGCTTCTGAAGTCACCACCAAAATGCTGTGCTGCTGCTGACCCACACAGTG--- 1180

Qy   939 CGGGCCCCAGGCAAACTTGCACGACGCCTGCCAGGGCAATTCGGCAGGGCGCCCCCTGGTGTG 998
Db  1181 -------------------GAAAACAGATTCCTGCCAGGGACACTTCAGGGGACCCCCTGTCTG 1225

Qy   999 TCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGGGCTGGGTGTGGACA 1058
Db  1226 TTCCCTCCAAGGCCGCATGACTTTGACTGGAATTGTGACTGGAATTGTGAGCTGGGGCGGTGGGATGTGCCCT 1285

Qy  1059 GAAGGAGATGTCCCGGGTGTGTACACAAAGGTTACCAACTACCTAGACTGATTCG 1112
Db  1286 GAAGGACAAGCCAGGCGTCTACACGAGAGTCTACACACTTCTTACCCTGGATCCG 1339
```

```
RESULT 13
US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US2002015139A1
; GENERAL INFORMATION:
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: McNeill, Patricia D.
;  APPLICANT: Retter, Marc W.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C15
;  CURRENT APPLICATION NUMBER: US/09/850,716A
;  CURRENT FILING DATE: 2001-05-07
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;  LENGTH: 1475
;  TYPE: DNA
;  ORGANISM: Homo sapien
US-09-850-716A-122

Query Match             15.1%;  Score 170.8;  DB 10;  Length 1475;
Best Local Similarity   51.3%;  Pred. No. 9.6e-44;
Matches  551;  Conservative    0;  Mismatches  457;  Indels  66;  Gaps  4;

Qy    84 CTGCTACTTTGGGAATGGTCAGCTACCGTGGCACGCACAGCCTCACCGAGTCGGGTGC 143
Db   287 CTGCTATGAGGGAATGGTCACTTTTACCGAGGAAAAGGCCAGCACTGACACCATGGGCCG 346

Qy   144 CTCCTGCCTCCCGTGGAATTCCATGATCCTGCACTGTCCTTCCTTCAGCAAACATAATTACTGC 203
Db   287 GCCCTGCCTGGAACTCTGCCACTGTCCTTCAGCAAACGTACCATGCCCACAGATC 346

Qy   204 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263
Db   347 GCCCTGCCTGGGAACTCTGGGGTGATCAGCGCCACACTGCCGCCACCAGATC 406

Qy   144 CTCCTGCCTCCCGTGGAATTCCATGATCCTGATGATCAGCAGCGCCCACACTGCTTCAGGA 203
Db   347 GCCCTGCCTGGGAACTCTGGGGTGATCAGCGCCACACTGCTGCTTCATTGA 406

Qy   347 GCCCTGCCTGGTGGGTGATCAGCGCCACACTGCCCACCATGCCCACCAGATC 406

Qy   204 CAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263

Qy   407 TGATGCTCTTCAGCTGGGGAATCAGCTGGGCAAACATAATTACTGCCGAGAACCCCAGAACCCGGAG 466
```

```
Query Match            15.1%;   Score 170.8;  DB 10;  Length 1475;
Best Local Similarity  51.3%;   Pred. No. 9.6e-44;
Matches  551;  Conservative  0;  Mismatches .457;  Indels  66;  .Gaps  4;
```

```
Qy    84  CTTGCTACTTTGGAATGGGTCAGCCTACCGTGGCACGCACAGCCTCACCGAGTCGGGTGC  143
              |||||  ||||  ||  ||||||  ||||||||  ||||||||| |||||||||| ||||||  ||
Db   287  CTTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAAGGCCAGCACTGACCACCATGGGCG  346

Qy   144  CTCCTGCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTTTACACAGCACAGAAACCC  203
Db   347  GCCTGCGTCCTGGAACTCTGGAACTCTGCCACTGTCCTTCAGCAAAGTACCACCACAGATC  406

Qy   204  CAGTGCCCAGGCACTGGGCTGGGCAAACATAATTACTGCCGGAATCCTGATGGGGATGC  263
Db   407  TGATGCTCTTCAGCTGGGCCTGGGGAAACATAATTACTGCAGGAACATAATTACTCAGGAG  466

Qy   264  CAAGCCCTGTGCCACGTGCTGCTGAAGAACCGCAGCTGACGTGGGAGTACTGTGATGTGCC  323
Db   467  GCGACCCTGTGCTATGTGCAGGTGGGCCTAAAGCCGCTTGTCCAAGGACTGCATGGTGCA  526

Qy   324  CTCCTGCTCCACCTGCGGGCCTGAG---------------  347
              ||||||  |||
Db   527  TGACTGCGCAGATGAGATGGAAAAAGCCCTCCTCCTCCCAGAAGAGGAGGCTCTTCGCGACATCGCCCTC  586

Qy   348  ---ACAGTACAGCCTCTGAGGCCCATCTTTGCCAAGGACTTGGGGGAAAATACAGT  404
Db   587  CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGGAGAATTCACCACCATCGAGAA  646

Qy   405  CCACCCCTGGCAGCTGCCATCTTGCCAAGCACAGAGGTCGCCCCGGAGAGCGGTTCCT  464
Db   647  CAGCCCTGGTTTGCGGCCATCTACAGGAGGACACCGGGG------GGGCTCTGTCACCTACGT  703

Qy   465  GTGCGGGGGCACTGGGCTGGGCAAATCGGATTCGTCCCCGTG  524
Db   704  GTGTGGAGGCAGCCTCATCCTGGGCGTGATCTTTGGGCAGAACATACCGGGTGGTCCCTGG  763

Qy   525  GAGGTTTCCGCCCCACCACCTGAGGCTGCCCCTTGGGTGGTCCCTGG  584
Db   764  TTACCCAAAGAAGGAGGACTACATCGTTCACCTGGGTGCCTCAAGGCTTAACTCCAACAC  823

Qy   585  CGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCATGATGA  644
Db   824  GCAAGGGGAGATGAAGTTTGAGGTGCTGCTGCTGACCCCACAGTG---  883

Qy   645  CA------CTTACGACAATGACATTGCCTGCTGCAGCTGAAATCGGATTCGTCCCGCTG  698
Db   884  CACGCTTGCTCACCACCACGACATTGCCTTGCTGAAGATCAGCAGCGCCACACACTGCTTCATTGA  943

Qy   699  TGCCCAGGAGAGCAGCGGTCGTCGTCCCACTGTGCCTTCCCCGCGGACCTGCAGCTGCC  758
Db   944  TGCGGCAGCCATCCCGACTATCCCGGAGTGTCAGCAGCC  1003

Qy   759  GGACTCGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGCCTTGTCTCTCCTTTCTA  818
Db  1004  GTTTGGCACAAGCTGTGAGATCACTGCCTTTGGAAAAGAGAATTCTACCGACTATCTCTA  1063

Qy   819  TTCGGAGGGGCTGAAGGAGGGCTCATGTGCAGCTGAGGCTCATGGTGCAGCAGTG  878
Db   819  TTCGGAGGGGCTGAAGGAGGGCTCATGTGCAGCTGAGGCTCATGGTGCAGCAGTG
```

```
RESULT 15
US-09-466-396A-122
; Sequence 122, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-122

Query Match         15.1%;   Score 170.8;   DB 11;   Length 1475;
Best Local Similarity   51.3%;
Matches  551;  Conservative   0;  Mismatches 457;  Indels  66;  Gaps   4;

Pred. No. 9.6e-44;
```

```
Qy     84 CTGCTACTTTGGGAATGGGTCAGCTCAGCCTCACCGTGCCACCGACCAGCTCGGGTGC 143
Db    287 CTGCTATGAGGGGAATTGGTCACTTTTACCGAGGAAAGGCCAGCACTGACACACCATGGGGCCG 346

Qy    144 CTCCTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTTACACAGCACAGAACCC 203
Db    347 GCCCCTGCCTGCCCTGGAACTCTGCCACTGTCCTTCCTGCCACTGTCCTGCCAAACGTACCATGCCCACACATC 406

Qy    204 CAGTGCCCAGGCACTGGGCCTGGCACAAACATAATTACTGCCGGAATCCTGATGGGGATGC 263
Db    407 TGATGCTCTTCAGCTGGGCTGGCCTGGGCCTGGGAAACATAATTACTGCAGGAAACCCAGACAACCGGAG 466

Qy    264 CAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCGTGACGTGGAGTACTGTGATGTGCC 323
Db    467 GCGACCCGGTGCTATGTGCAGGTGGCCTAAAGCCCGGCTTGTCCAAGAGTGCATGGTGCA 526

Qy    324 CTCCTGCTCCACCTGCGCGGCCTGAG----------------------------------- 347
Db    527 TGACTGCGCAGATGAGGGGAATGCCGGGAAAAAAGCCCCTCCTCCTCCAGAAGAATTAAAATTTCAGTGTGG 586

Qy    348 ---ACAGTACAGCCAGCCTCAGTTTCGCATCAAAAGAGGGGCTCTTCGCCCGACATCGCCTC 404
Db    587 CCAAAAGAACTCTGAGGCCCCGCTTTAAGATTATTGGGGGAGAATTCACCACCATCGAGAA 646

Qy    405 CCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGAGGGTCGCCCGGAGAGCGGTTCCT 464
Db    647 CCAGCCCTGGTTTGCGGCCATCTACAGGAGGCCACCGGGG--GGGCTCTGTCACCTACGT 703

Qy    465 GTGCGGGGGCATACTCATCAGCTCCTGCTGCGATTCTCTCGCCCCACTGCTTCCAGGA 524
Db    704 GTGTGGAGGCAGCCTCATCAGCCCTTGCTGGGTGATCAGCGCCACACACTGCTTCATTGA 763

Qy    525 GAGGTTTCCGCCCCACCTGACGGTGATCTTGGGCAGCAGAACATACGGGTGGTCCCTGG 584
Db    764 TTACCCAAAGAAGGAGGACTACATCGTCTACCTGGGTCGCTCAAGGCTTAACTCCAACAC 823

Qy    585 CGAGGAGCAGCAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGA 644
Db    824 GCAAGGGGGAGGATGAAGTTTGAGGTGGAAAACCTCATCCTACACAAGGACTACAGCGCTGA 883

Qy    645 CA------CTTACGACAATGACATTGCCTGCTGCAGCTGAAATCGGATTCGTCCGCTG 698
Db    884 CACGCTTGCTCACCACAACGACATTGCCTTGCTGAAGATCCGTTCCAAGGACGGGCAGGTG 943

Qy    699 TGCCCAGGAGGAGCGGTGGTCCGCACTGTCGGTCCTTCCCCGGCGGGGACCTGCAGCTGCC 758
Db    944 TGGCGACGCATCCCGGACTATACAGACCATCTGCCTGCCCTCGATGTATAACGATCCCCA 1003

Qy    759 GGACTGGACGGAGTGTGAGCTCTCCGGTACGGCAAGCATGAGGCCTTGTCTCCTTCTA 818
Db   1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAGAGAATTCTACCGACTATCTCTA 1063

Qy    819 TTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACA 878
Db   1064 TCCGGAGCAGCTGAAGATGACTGCTGTGTGAAGCTGATTCCCACCGGGAGTGTCAGCAGCC 1123

Qy    879 ACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTCGTGGAGACACTCGGAGCGG 938
Db   1124 CCACTACTACGGCTTCTGAAGTCACCACCAAAATGCTGTGTGCTGTGCTGACCCACAGTG--- 1180

Qy    939 CGGGCCCCAGGCAAACTTGCACGACGCGCTGCCAGGGCGATTCGGAGGGCCCCCTGGTGTG 998
Db   1181 ------------GAAAACAGATTCCTGCCAGGAGACTCAGGGGGACCCCCGTCTG 1225

Qy    999 TCTGAACGATGGCCGCATGACTTTGGTGGCATCATCAGCTGGGCGTGTGGACA 1058
Db   1226 TTCCCTCCAAGGCCGCATGAGACTTTGACTGGGAATTGTGAGCTGGGCCGTGGATGTGCCCT 1285

Qy   1059 GAAGGATGTCCCGGTGTGTACACAAAGGTTACCAACTACCTGACTGATTCG 1112
Db   1286 GAAGGACAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCCTGATCCG 1339
```

Search completed: August 19, 2003, 14:22:22
Job time : 321.307 secs

OM nucleic - nucleic search, using sw model

```
            Copyright   GenCore version 5.1.6
            Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: August 18, 2003, 22:44:33 ; Search time 3299.95 Seconds
(without alignments)
8307.845 Million cell updates/sec

Title: US-09-987-455-2
Perfect score: 1128
Sequence: 1 atgaaaagacagctatcgc.........ttcgtgacaacatgcgaccg 1128

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
         EST:*
          1: em_estba:*
          2: em_esthum:*
          3: em_estin:*
          4: em_estmu:*
          5: em_estcv:*
          6: em_estpl:*
          7: em_estro:*
          8: em_htc:*
          9: gb_est1:*
         10: gb_est2:*
         11: gb_htc:*
         12: gb_est3:*
         13: gb_est4:*
         14: gb_est5:*
         15: em_estfun:*
         16: em_estom:*
         17: em_gss_hum:*
         18: em_gss_inv:*
         19: em_gss_pln:*
         20: em_gss_vrt:*
         21: em_gss_fun:*
         22: em_gss_mam:*
         23: em_gss_mus:*
         24: em_gss_pro:*
         25: em_gss_rod:*
         26: em_gss_phg:*
         27: em_gss_vrl:*
         28: gb_gss1:*
         29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	895	79.3	930	13	BX329047	BX329047 BX329047
2	853	75.6	912	13	BU149958	BU149958 AGENCOURT
3	850.2	75.4	962	13	BX389609	BX389609 AGENCOURT
4	818.6	72.6	891	13	BU146167	BU146167 AGENCOURT
5	813.8	72.1	993	13	BQ278177	BQ278177 AGENCOURT
6	811.2	71.9	883	13	BQ690978	BQ690978 AGENCOURT
7	807.6	71.6	918	13	BX389608	BX389608 AGENCOURT
8	796.2	70.6	930	13	BU845014	BU845014 AGENCOURT
9	782.8	69.4	888	13	BU157720	BU157720 AGENCOURT
10	765.2	67.8	950	13	BQ927899	BQ927899 AGENCOURT
11	764.4	67.8	916	12	BT765247	BT765247 603050338
12	706.8	62.7	948	10	BE616613	BE616613 601278864
13	700.4	62.1	958	12	BG763582	BG763582 602735987
14	694.4	61.6	974	13	BQ960216	BQ960216 AGENCOURT
15	693.4	61.5	709	10	BE732704	BE732704 601571366
16	688.2	61.0	1060	12	BM566999	BM566999 AGENCOURT
17	677	60.0	1042	10	BE732413	BE732413 601569644
18	672	59.6	957	13	BQ687538	BQ687538 AGENCOURT
19	665.8	59.0	948	13	BQ687779	BQ687779 AGENCOURT
20	665.2	59.0	704	9	AU134301	AU134301 AU134301
21	655.4	58.1	818	13	AU124602	AU124602 AU124602
22	655.2	58.1	903	13	BQ689095	BQ689095 AGENCOURT
23	653.2	57.9	830	13	BQ896938	BQ896938 AGENCOURT
24	653.2	57.8	867	13	BU179903	BU179903 AGENCOURT
25	652.2	57.8	904	13	BQ689840	BQ689840 AGENCOURT
26	652.2	57.8	911	13	BQ684734	BQ684734 AGENCOURT
27	651	57.7	818	12	BM006474	BM006474 603615409
28	649.6	57.6	797	13	BQ879911	BQ879911 AGENCOURT
29	642.2	56.9	1037	13	BU157493	BU157493 AGENCOURT
30	631.2	56.0	813	10	BG120780	BG120780 602348540
31	629.8	55.8	663	9	AU134463	AU134463 AU134463
32	624	55.3	624	10	BE549489	BE549489 601278502
33	585.4	51.9	587	9	BE394944	BE394944 601308225
34	561	49.7	561	9	AL047284	AL047284 DKF2p586H
35	553.4	49.1	555	14	CB128855	CB128855 K-EST0178
36	535.4	47.5	537	14	CB155052	CB155052 K-EST0213
37	534	47.2	964	12	BG769384	BG769384 602742678
38	520.4	46.1	905	13	BU859271	BU859271 AGENCOURT
39	514.8	45.6	883	14	CD359959	CD359959 AGENCOURT
40	514.4	45.6	516	14	CB140923	CB140923 K-EST0194
41	493.4	43.7	916	13	BU157241	BU157241 AGENCOURT
42	487.8	43.2	527	9	AW630119	AW630119 hh78e06.y
43	482.8	42.8	553	9	AW394127	AW394127 MR2-TT001
44	474.8	42.1	917	13	BU190904	BU190904 AGENCOURT
45	472.8	41.9	937	13	BQ679501	BQ679501 AGENCOURT

ALIGNMENTS

```
RESULT 1
BX329047/c
LOCUS       BX329047               930 bp    mRNA    linear   EST 01-MAY-2003
DEFINITION  BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 3-PRIME, mRNA sequence.
ACCESSION   BX329047
VERSION     BX329047.1  GI:30307809
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4498.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAI021ZD08_C
            S01980 1&cluster=4498.r. Contact : Feng Liang Email :
            fliang@lifetech.com URL : http://fulllength.invitrogen.com/
            InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID :
```

```
FEATURES             CS0BAI02IZD08_CS01980_1
     source          1..930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI039YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT       184 a      268 c      282 g      194 t      2 others
ORIGIN
```

```
Qy     67 TCTGAGGGAAACAGTGACTGACTGCTACTTTGGGAATGGTCAGCCTACCGTGGCACGCACAGC  126
          | |||||||||||||  ||||||||||||   |||||||||||||||||  ||||||||||||||||
Db    927 TTTGAGGGAAACAGTGACTGACTGCTAC-TTGGGAATGGTCAGCCTACGTGGGCACGGACAGC  869

Qy    127 CTCACCGAGTCGGGTGCCTCCGTGCCTCCCGTGGAATTCCATGATCCTGATAGGCAAGGTT  186
          ||||||||||||||||||||||||||  |  ||||||||||||| || | ||||||||||||||||
Db    868 CTCACCGAGTCGGGTGCTNCCGTGCCTCCCGTGAATTCCATGATCCTGATAGGCAAGGTT  809

Qy    187 TACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGG  246
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 TACACAGCACAGAACNCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCGG  749

Qy    247 AATCCTGATGGGATGCCAAGCCCTGGTTGCCACGTGCCACCGGCAGCCTGACGTGG  306
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    748 AATCCTGATGGGATGCCAAGCCCTGGTTGCCACGTGCCACCGGCAGCCTGACGTGG  689

Qy    307 GAGTACTGTGATGTGCCCCTGCTCCACCTGCGCCGCCTGAGACAGTACAGCCAGCCTCAG  366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 GAGTACTGTGATGTGCCCCTGCTCCACCTGCGCCGCCTGAGACAGTACAGCCAGCCTCAG  629

Qy    367 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC  426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 TTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCCTGGCAGGCTGCCATC  569

Qy    427 TTTGCCAAGCACAGGAGGTGCCCCGGAGACGCGGTTCCTGTGCGGGGGCATACTCATCAGC  486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568 TTTGCCAAGCACAGGAGGTGCCCCGGAGACGCGGTTCCTGTGCGGGGGCATACTCATCAGC  509

Qy    487 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCGCCCCACCTG  546
          |||||||||||||||||||||||||||||||||||||||||||
Db    508 TCCTGCTGGATTCTCTCTGCCGCCACTGCTTCCGCCCCACCTG  449

Qy    547 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGGAGGAGCAGAAATTTGAA  606
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 ACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGGAGGAGCAGAAATTTGAA  389

Qy    607 GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG  666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 GTCGAAAAATACATTGTCCATAAGGAATTGATGATGACACTTACGACAATGACATTGCG  329

Qy    667 CTGCTGCAGCTGAAATCGGATTCTCCCGTGTGCCCAGGAGAGCAGCGTGGTCCGCACT  726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 CTGCTGCAGCTGAAATCGGATTCTCCCGTGTGCCCAGGAGAGCAGCGTGGTCCGCACT  269

Qy    727 GTGTGCCTTCCCCCGGCCGACCTTGCAGCTGCCGACTGCCGAGTGTGAGCTCTCCGGC  786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 GTGTGCCTTCCCCCGGCCGACCTTGCAGCTGCCGACTGCCGAGTGTGAGCTCTCCGGC  209

Qy    787 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  846
          |||||||||||||||||||||||  ||  ||||||||||||||||||||||||||||||||
Db    208 TACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC  149

Qy    847 AGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACAGTCACCGAC  906
```

RESULT 2
```
LOCUS       BU149958                 912 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8137213 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184119 5', mRNA sequence.
ACCESSION   BU149958
VERSION     BU149958.1  GI:22663490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13572 row: m column: 16
            High quality sequence start: 18
            High quality sequence stop: 644.
FEATURES             Location/Qualifiers
     source          1..912
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6184119"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site 2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT       189 a      272 c      271 g      179 t      1 others
ORIGIN
```

```
Query Match     75.6%;  Score 853;  DB 13;  Length 912;
Best Local Similarity  98.9%;  Pred. No. 2.5e-211;
Matches  901;  Conservative    0;  Mismatches    5;  Indels    5;  Gaps    4;

Qy    197 AGAACCCCAGTGCCCAGGCACTGG--CCTGGCCAAACATAATCATTACTGCGGCAAGAATCCTGA  254
          |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db      1 AGAACCCCAGTGCCCAGGCACTGGGGCCTGGGCCAAACATAATTACTGCGGCAAGAATCCTGA  60

Qy    255 TGGGGGATGCCAAGCCCTGTGCCACGTGCCACGTGCTGAAGAACCGCAGCGCTGACGTGGGAGTACTG  314
```

AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4498.r For more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAF025ZG09_AF02376_3&cluster=4498.r.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09_AF02376_3.
FEATURES   Location/Qualifiers
     source   1..962
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DIO39YK15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT   203 a   290 c   277 g   191 t   1 others
ORIGIN

Query Match   75.4%; Score 850.2; DB 13; Length 962;
Best Local Similarity 96.6%; Pred. No. 1.4e-210;
Matches 879; Conservative   0; Mismatches 29; Indels 2; Gaps 1;

```
Qy    208 GCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGAATCCTGATGGGGATGCCAAG 267
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCCAGGCACTGGNGCCTGGGCAAACTAATTACTGCCGAATCCTGATGGGGATGCCAAG 60

Qy    268 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGGGAGTACTGTGATGTGCCCTCC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGTGCCCTCC 120

Qy    328 TGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGGCTC 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGGCTC 180

Qy    388 TTCGCCGACATCGCCTCCCACCCCTGGCGAGGCTGCCATCTTGCCAAGCACAGGAGGTCG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCGCCGACATCGCCTCCCACCCCTGGCGAGGCTGCCATCTTGCCAAGCACAGGAGGTCG 240

Qy    448 CCCGGAGAGCGGTTCCTGTGCGGGGGCATACTACCATCAGCCTCCTGGATTCCTCTGCC 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCGGAGAGCGGTTCCTGTGCGGGGGCATACTACCATCAGCCTCCTGGATTCCTCTGCC 300

Qy    508 GCCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACA 567
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACA 360

Qy    568 TACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTGTCCAT 627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTGTCCAT 420

Qy    628 AAGGAATTCGATGATGACACTTACGACAATGACATTGCGTCGTGCTGCTGCAGTGAAATCGGAT 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAGGAATTCGATGATGACACTTACGACAATGACATTGCGTCGTGCTGCTGCAGTGAAATCGGAT 480

Qy    688 TCGTCCCGTGTGCCCCAGAGAGCAGCAGCAGCGTGGTCCGACACTGTGTGCCTTCCCCCGGCGGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCGTCCCGTGTGCCCCAGAGAGCAGCAGCAGCGTGGTCCGACACTGTGTGCCTTCCCCCCGGCGGAC 540

Qy    748 CTGCAGCTGCCGACTGCGACGAGTGTGAGCTCTCCGGCTACGGCCAAGCATGAGGGCCTTG 807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGCAGCTGCCGACTGCGACGAGTGTGAGCTCTCCGGCTACGGCCAAGCATGAGGGCCTTG 600
```

```
Db     61 TGGGGATGCCAAGCCCTGTGCCACGTGCCACCTGCTGCTGAAGAACCGCAGGCTGGGAGTACTGTG 120

Qy    315 TGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCAT 374
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGATGTGCCCTCCTGCTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTTCGCAT 180

Qy    375 CAAAGGAGGGGCTCGCCCGGAGAGGTCGCCCGCCATCTCCACCCCCTGGCAGGCTGCCATCTTGCCAA 434
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAAAGGAGGGGCTCGCCCGGAGAGGTCGCCCGCCATCTCCACCCCCTGGCAGGCTGCCATCTTGCCAA 240

Qy    435 GCACAGGAGGTCGCCCGGAGAGGGCATACTCATCAGCTCCTGCTG 494
          |||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCACAGGAGGTCGCCCGGAGAGGGCATACTCATCAGCTCCTGCTG 300

Qy    495 GATTCTCTGCCGCCGCCACTGCTTCCAGGAGGAGGTTCCGCCCCACCACCTGACGGTGAT 554
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GATTCTCTGCCGCCGCCCACTGCTTCCAGGAGGAGGTTTCCGCCCCACCACCTGACGGTGAT 360

Qy    555 CTTGGGCAGAACATACCGGGTGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAA 614
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTTGGGCAGAACATACCGGGTGGTGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAA 420

Qy    615 ATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCCGTCGTGCA 674
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCCGCTGCTGCA 480

Qy    675 GCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCT 734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GCTGAAATCGGATTCGTCCCGTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCCT 540

Qy    735 TCCCCCCGGCGGACACTGCAGCTGCCGACTGCGACGAGTGTGAGCTCTCCGGCTACGGCCAA 794
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCCCCCCGGCGGACACTGCAGCTGCCGACTGCGACGAGTGTGAGCTCTCCGGCTACGGCCAA 600

Qy    795 GCATGAGGGCCTTGCTCCTTTCTATTCGGACGCGCTGAAGGAGCTCATGTCAGACTGTA 854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCATGAGGGCCTTGTCTCCTTTCTATTCGGACGCGCTGAAGGAGCTCATGTCAGACTGTA 660

Qy    855 CCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCT 914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCT 720

Qy    915 GTGTGCTGGAGACACTCGGAGCGGCGGGCCCAGGCAAACTTGCACGACGCCTGCCAGGG 974
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTGTGCTGGAGACACTCGGAGCGGCGGGCCCAGGCAAACTTGCACGACGCCTGCCAGGG 780

Qy    975 -CGATTCGGGAG-CCCCCTGGTGTCTGAACGAGATGGCCGATGACGTTGGTGGGGCATC 1032
          ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db    781 NCGATTCGGGAGGCCCCCTGGTGTCTGAACGAGATGGCCGATGACGTTGGTGGGGCATC 840

Qy   1033 ATCAGCTGGGGCTGGGCTGTGACGAAGGATGT-CCCGGGTGTGTACACAAAGGTTAC 1091
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    841 ATCAGCTGGGGCTGGGCTGTGACGAAGGATGTCCCCGGGTGTGTACACCAAGGTTAC 900

Qy   1092 CAACTACCTAG 1102
          |||| ||||||
Db    901 AAACTACCTAG 911
```

RESULT 3
BX389609
LOCUS   BX389609   962 bp   mRNA   linear   EST 08-MAY-2003
DEFINITION BX389609 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DIO39YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389609
VERSION   BX389609.1 GI:30463470
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 962)

```
BASE COUNT      178 a      266 c      271 g      176 t
ORIGIN

Qy          CCCAGTGCCCAGGCACTGGGCTCTGGGATGGCTGAGGAGGGCTCATGTCAGACTGTACCCATCCAGCGGC    867
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          CCCAGTGCCCAGGCACTGGGCTCTGGGATGGCTGAGGAGGGCTCATGTCAGACTGTACCCATCCAGCCCC    660
```

Query Match   72.6%;   Score 818.6; DB 13;   Length 891;
Best Local Similarity 98.2%;   Pred. No. 2.3e-202;
Matches 849;   Conservative 0;   Mismatches 14;   Indels 2;   Gaps 2;

```
Qy      202  CCCAGTGCCCAGGCACTGGGCTCTGGGATGGCTGAGGAGGGCTCATGTCAGACTGTACCCATCCAGCGGC  261
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1  ...CCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTTGCCGAAATCCTGATGGGGAT         60

Qy      262  GCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCTGGAGTGGGAGTACTGTGATGTG             321
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61  GCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGCTGGAGTGGGAGTACTGTGATGTG             120

Qy      322  CCCTCCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGA            381
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121  CCCTCCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGA            180

Qy      382  GGGCTCTTGCTCCGGCCGCTCCCACCCCTGCCAGCCTGCCATCTTTGCCAAGCACAGG     441
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181  GGGCTCTTTGCTCCGGCCGCTCCCACCCCTGCCAGCCTGCCATCTTTGCCAAGCACAGG     240

Qy      442  AGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGCTCCTGCTCGGATTCTC    501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241  AGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGCTCCTGCTCGGATTCTC    300

Qy      502  TCTGCGCCACTGCTTCCAGAGAGAGGTTCCGCCCCACCACTGACCGGTGATCTTGGGC      561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301  TCTGCGCCACTGCTTCCAGAGAGAGGTTCCGCCCCACCACCTGACCGGTGATCTTGGGC     360

Qy      562  AGAACATACCGGTGGTCCTGCCGAGGAGGAGCAGAAATTGAAGTCGAAAAATACATT      621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361  AGAACATACCGGTGGTCCTGCCGAGGAGGAGCAGAAATTGAAGTCGAAAAATACATT      420

Qy      622  GTCCATAAGGAATTCGATGATGACACTTACGACAATGACAATTGGCTGCTGCAGCTGAAA   681
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421  CTTCCATAAGGAATTCGATGATGACACTTACGACAATGACAATTGGCTGCTGCAGCTGAAA   480

Qy      682  TCGGATTCGTCCCGCTGTCCCCGTGTCCCAGGAGAGCAGCGGTGGTCCGCACTGTGTGCCTTCCCCCG   741
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481  TCGGATTCGTCCCGCTGTCCCCGTGTCCCAGGAGAGCAGCGGTGGTCCGCACTGTGTGCCTTCCCCCG   540

Qy      742  GCGGACCTGCAGCTGCCCGACTGACGAGTGTGAGCTCTCCGGCTACGGCAAGCCATGAG   801
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541  GCGGACCTGCAGCTGCCCGACTGACGAGTGTGAGCTCTCCGGCTACGGCAAGCCATGAG   600

Qy      802  GCCTGTGTCCTTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCC   861
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db      601  GCCTTGTGTCCTTTCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCC   660

Qy      862  AGCGCCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCT   921
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661  AGCGCCTGCACATCACACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCT   720

Qy      868  TGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGGGAGAC    927
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661  TGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGGGAGAC    720

Qy      922  GGAGACACTCGGACGCGGCGGC-GGGCCCCAGGCAAACTTGCACGAGGCCTGCCAGGGCGGATTC   980
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      721  GGAGACACTCGGACGCGGCGGGCCCCAGGCAAACTTGCACGAGACGCGGCTGCCAGGGCGATTC    780

Qy      928  ACTCGGAGCGGCGGGCGGCCCCCAGGCAAACTTGCACGAGACGCGCTGCCAGGGCCGATTCCGGAGGC   987
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721  ACTCGGAGCGGCGGCGGCCCCAGGCAAACTTGCACGAGACGCGCTGCCAGGGCCGATTCGGGAGGC    780

Qy      981  GGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCAT-GACTTTGTGGGCATCATCAGCT  1039
             ||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      781  GGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGTGGGCATCATCAGCT   840

Qy      988  CCCCTGGTGTGTCTGAACGATGGCCGCATGACTTTGGTGGGCATCATCAGCTGGTGTACACAAAGGTTGGGGCCTG  1047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781  CCCCTGGTGTGTCTGAACGATGGCGCGCATGACTTTGGTGGGCATCATCAGCCTGGGGCCTG  840

Qy     1048  GGCTGGTGGACAGAAGGATGTGCCCCGGGTGTGTACACAAAGGTTACCAACCTAGACTAGG  1107
             ||||||||| ||||||||||||||||||||||||||| ||| |||||
Db      841  GGCTGTGGAACAGAAGATGTGCCCGGGTTGTTCCCACCCAG--TTACCACTTACTAAACTGG  898

Qy     1108  ATTCGTGACA  1117
             ||||| |||
Db      899  ATTCGGGACA  908
```

RESULT 4
BU146167
LOCUS   BU146167   891 bp   mRNA   linear   EST 03-SEP-2002
DEFINITION AGENCOURT_7981262 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6185661 5', mRNA sequence.
ACCESSION   BU146167
VERSION   BU146167.1   GI:22659699
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 891)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13576 row: m column: 22
            High quality sequence stop: 665.
FEATURES   Location/Qualifiers
     source   1..891
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6185661"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGGTCCG-3' and
                     5'-GACTAGTTCTAGATCGGCAGCGGCCGCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life

RESULT 5
BQ278177
LOCUS   BQ278177   993 bp   mRNA   linear   EST 07-MAY-2002
DEFINITION AGENCOURT_7061709 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5805153

```
Db    360  CATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGT  419
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Qy    720  CCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCGGACTGTGACGGAGTGTGAGCT  779
                ││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    420  CCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCGGACTGTGACGGAGTGTGAGCT  479

Qy    780  CTCCGGCTACGGCAAGCATGAGGCCTTGCTCCTCCTTTCATTCGGAGCGGCTGAAGGAGGC  839
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    480  CTCCGGCTACGGCAAGCATGAGGCCTTGCTCCTCCTTTCATTCGGAGCGGCTGAAGGAGGC  539

Qy    840  TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGT  899
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    540  TCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGT  599

Qy    900  CACCGACACAACATGCTGTGCTGAGACACTCGGAGCGCGCGGCGCCCCAGGCAAACTTGCA  959
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    600  CACCGACACAACATGCTGTGCTGAGACACTCGGAGCGCGCGGCGCCCCAGGCAAACTTGCA  659

Qy    960  CGACGCCTGCCAGGGGATTCGGGAGGGGCCCCCCGTGGTGTCTGAACGATGGCCGCATGAC  1019
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    660  CGACGCCTGCCAGGGGATTCGGGAGGGGCCCCCCGTGGTGTCTGAACGATGGCCGCATGAC  719

Qy   1020  TTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTA  1079
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    720  TTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAAGGATGTCCCGGGTGTGTA  779

Qy   1080  CACAAAGGTTACCAACTACCTAGACTGGATTCGTGACACACATGCGGACCG  1128
                ││││││││││││││││││││││││││││││││││││││││││││││││││
Db    780  CACCAAGGTTACCAACTACCTAGACTGGATTCGTGACAACATGCGGACCG  828
```

RESULT 6
```
BQ690978
LOCUS       BQ690978                 883 bp    mRNA            linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
            5', mRNA sequence.
ACCESSION   BQ690978
VERSION     BQ690978.1  GI:21816294
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2367 row: o column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1..883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6209120"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
```

```
ACCESSION   BQ278177
VERSION     BQ278177.1  GI:20488385
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 993)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2043 row: o column: 10
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1..993
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5805153"
                     /tissue_type="teratocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC_Library."
BASE COUNT      235 a    295 c    269 g    194 t
ORIGIN
```

```
Query Match             72.1%;  Score 813.8;  DB 13;  Length 993;
Best Local Similarity   99.6%;  Pred. No. 4.3e-201;
Matches  826;  Conservative  0;  Mismatches  2;  Indels  1;  Gaps  1;

Qy    300  GACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCTGCGAGACAGTACAGCCA  359
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db      1  GACGTGGGAGTACTGTGATGTGCCCTCCTGCT-CACCTGCGGCTGCGAGACAGTACAGCCA   59

Qy    360  GCCTCCAGTTCGCATCAAAGGAGGGGCCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGC  419
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db     60  GCCTCCAGTTCGCATCAAAGGAGGGGCCTCTTCGCCGACATCGCCTCCACCCCTGGCAGGC  119

Qy    420  TGGCCATCTTTGCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCCTGTCGCGGGGCATACT  479
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    120  TGGCCATCTTTGCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCCTGTCGCGGGGCATACT  179

Qy    480  CATCAGCTCCCTGGATTCTCTCGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCA  539
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    180  CATCAGCTCCTGCTGGATTCTCTCGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCA  239

Qy    540  CCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCTCCTGGCGAGGAGGAGCAGAA  599
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    240  CCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGCTCCTGGCGAGGAGGAGCAGAA  299

Qy    600  ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA  659
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││││
Db    300  ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCCATGATGACACTTACCACACAATGA  359

Qy    660  CATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCCAGGAGAGCAGCGTGGT  719
                │││││││││││││││││││││││││││││││││││││││││││││││││││││││││
```

```
                    (Stratagene) and Superscript II RT (Life Technologies).
                    Note: this is a NIH_MGC Library."

BASE COUNT     184 a    265 c    259 g    175 t    0;
ORIGIN

Query Match     71.9%;    Score 811.2;  DB 13;  Length 883;
Best Local Similarity   99.0%;  Pred. No. 1.9e-200;
Matches  816;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy  178 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAAT 237
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GGCAAGGTTTACACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAAT 60

Qy  238 TACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 297
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy  298 CTGACGTGGGAGTACTGTGATGTGCCCCTCGCTGCGGCCTGAGACAGTACAGC 357
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGACGTGGGAGTACTGTGATGTGCCCCTCGCTGCGGCCTGAGACAGTACAGC 180

Qy  358 CAGCCTCAGTTTCGCATCAAAGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAG 417
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CAGCCTCAGTTTCGCATCAAAGAGGGGCTCTTCGCCGACATCGCCTCCCACCCCTGCCAG 240

Qy  418 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATA 477
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCATA 300

Qy  478 CTCATCAGCTCCTGGATTCTCTCTGCCACTGGGGTCCTGGGCAAACTTG 537
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CTCATCAGCTCCTGGATTCTCTCTGCCACTGGGGTCCTGGGCAAACTTG 360

Qy  538 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAG 597
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAG 420

Qy  598 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 657
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACACAT 480

Qy  658 GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTCCCAGGAGCAGCGGTG 717
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTCCCAGGAGCAGCGGTG 540

Qy  718 GTCCGCACTGTGTGCCTTCCCCCGGGACCTGCAGCTGCGGGACTGGACGTGTGAG 777
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GTCCGCACTGTGTGCCTTCCCCCGGGACCTGCAGCTGCGGGACTGGACGTGTGAG 600

Qy  778 CTCTCCGGCTACGGCAAGCATGAGGCCTTGCTCTTTGTCTTTTCTATTCGGAGCGGCTGAGAGGAG 837
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTCTCCGGCTACGGCAAGCATGAGGCCTTGCTCTTTGTCTTTTCTATTCGGAGCGGCTGAGAGGAG 660

Qy  838 GCTCCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACA 897
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GCTCCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACA 720

Qy  898 GTCACCGACAACATGCTGTGTGCCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTG 957
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GTCACCGACAACATGCTGTGTGCCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTG 780

Qy  958 CACGACGCCTGCCAGGGCGATTCGGGGACGGCCCCCTGGTGTGTCT 1001
        |||||||||||||||||||||||||||||||||||||||||||||
Db  781 CACGACGCCTGCCAGGGCGGAATTCGGGGAAGAGGCCCCGTGTGTGT 824
```

RESULT 7
BX389608
LOCUS   BX389608   918 bp   mRNA   linear   EST 08-MAY-2003
DEFINITION BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI039YK15 5-PRIME, mRNA sequence.
ACCESSION   BX389608
VERSION   BX389608.1 GI:30463469

```
FEATURES    Location/Qualifiers
source      1..918
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DI039YK15"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and cloned into the Not I and EcoR V
            sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT     189 a    273 c    264 g    186 t    6 others
ORIGIN

Query Match     71.6%;    Score 807.6;  DB 13;  Length 918;
Best Local Similarity   96.4%;  Pred. No. 1.7e-199;
Matches  855;  Conservative  0;  Mismatches  29;  Indels  3;  Gaps  3;

Qy  236 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGCTGAAGAACCGCA 295
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ATTACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCTGCTGAAGAACCGCA 76

Qy  296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTGCGGCCTGAGACAGTACA 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 GGCTGACGTGGGAGTACTGTGATGTGCCCTCGCTGCGGCCTGAGACAGTACA 136

Qy  356 GCCAGCCTCAGTTTCGCATCAAAGAGGGGCTCTTCGCCGACATCGCCTCCCCCTGC 415
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 GCCAGCCTCAGTTTCGCATCAAAGAGGGGCTCTTCGCCGACATCGCCTCCCCCTGC 196

Qy  416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA 475
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  197 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCCGGTTCCTGTGCGGGGGCA 256

Qy  476 TACTCATCAGCTCCTGGATTCTCTCTGCCACATCACAACATTTACTTAACAGAACA 535
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 TACTCATCAGCTCCTGGATTCTCTCTGCCACATCACAACATTTACTTAACAGAACA 316

Qy  536 CCCACCACCTGACGGTGATCTTGGGCAGAACATACGGGTGTCCTGGCGGAGGAGC 595
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 CCCACCACCTGACGGTGATCTTGGGCAGAACATACGGGTGTCCTGGCGGAGGAGC 376

Qy  596 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  377 AGAAATTTGAAGTCGAAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 436

Qy  656 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCAGGAGCAGCAGCG 715
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  437 ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCAGGAGCAGCAGCG 496

Qy  716 TGGTCCGGACTGTGTGCCTTCCCCCGGCGGACTGGACGTCAGCTGCGGGACTGGACGGAGTGTG 775
```

```
ORIGIN

Query Match        70.6%;    Score 796.2;   DB 13;   Length 930;
Best Local Similarity   99.4%;    Pred. No. 1.6e-196;
Matches  820;   Conservative    0;   Mismatches    3;   Indels    3;   Gaps    2;

Qy   300 GAGTGGGAGTACTGTGATGTGCCTCCCTGCTCCACCTGGCTCGGCGCTGAGACAGTACAGCCA 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GACGTGGGAGTACTGTGATGTGCCTCCTCCTGCTCCTCCACCTCGCGCGGCTGAGACAGTACAGCCA 60

Qy   360 GCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCTGGCGAGGC 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCCACCCTGGCGAGGC 120

Qy   420 TGCCATCTTTGCCAAGCACAGAGGTCGCCGGAGGAGTGGTTCCTGTGGGGGGCATACT 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCCATCTTTGCCAAGCACAGAGGTCGCCGGAGGAGTGGTTCCTGTGGGGGGCATACT 180

Qy   480 CATCAGCTCCTGCTGGATTCTCTCGCCGCCCACTGCTTCCAGGAGAGGTTCCGCCCCA 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CATCAGCTCCTGCTGGATTCTCTCGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCA 240

Qy   540 CCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGGGGGAGGAGAGCAGAA 599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCACCTGACGGTGATCTTGGGCAGAACATACCGGTGGTCCCTGGGGGGAGGAGAGCAGAA 300

Qy   600 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 360

Qy   660 CATTGCCGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCAGGAGAGCAGCGTGGT 719
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CATTGCCGCTGCAGCTGAAATCGGATTCGTCCCGCGTGTGCCCAGGAGAGCAGCGTGGT 420

Qy   720 CCGCACTGTGTGCCTTCCCCCGGGACCTGCAGCTGCCGGAACTGGACGGAGTGTGAGCT 779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCGCACTGTGTGCCTTCCCCCGGGACCTGCAGCTGCCGGAACTGGACGGAGTGTGAGCT 480

Qy   780 CTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGC 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGGAGGC 540

Qy   840 TCATGTCAGACTGTACCCATCAGCCGCTGACATCACAACATTACTTAACAGAACAGT 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCATGTCAGACTGTACCCATCAGCCGCTGACATCACAACATTACTTAACAGAACAGT 600

Qy   900 CACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGCGGCGCCCCAGGCAAACTTGCA 959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGCGGCGCCCCAGGCAAACTTGCA 660

Qy   960 CGACGCCTGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGAC 1019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGACGCCTGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCATGAC 720

Qy  1020 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAAGGATGTCCC-GGGTGT 1078
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   721 TTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACAGAAGGATGTCCCGGGGTGT 780

Qy  1079 ACA-CAAAGGTTACCAACTACCTTAGACTTGGATTCGTGACAACATG 1122
           ||| |||||||||||||||||||||||||||||||||||||||||||
Db   781 ACACCCAAGGGTACCAACTACCTTAGACTTGGATTCGTGACAACATG 825

Db   497 TGGTCCGACTGTGCCTTCCCCGGGACCTGCCGGACTGGACGGAGTCGTG 556
           |||||||||||||||||||||||||||||||||||||||||||||||||
Qy   776 AGGCTCTCCGGCTACGGCAAGCATGAGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGG 835
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 AGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCCTTTCTATTCGGAGGCGGCTGAAGG 616

Qy   836 AGGCTCATGTCAGACTGTACCCATCAGCCGCTGACATCACAACATTACTTAACAGAA 895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617 AGGCTCATGTCAGACTGTACCCATCAGCCGCTGACATCACAACATTACTTAACACANGA 676

Qy   896 CAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGCGGCGCCCCAGGCAAACT 955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   677 CAGTCACCGANCAACATGCTGTGTGCTGAGACACTCGNAGCGCGGCGCCCCA-GCAAACT 735

Qy   956 TGCACGACGCCTGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTCTGAACGATGGCCGCA 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   736 TGCACGACGCCTGCCAGGGGATTCGGGAGGCCCCCTGGTGTGTCTTGACGATTGCCCAA 795

Qy  1016 TGA-CTTTGGTGGGCATCATCAGCTGGGG--CCTGGGCGTGTGGACAGAAGGATGTCCCGGG 1073
           ||| |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db   796 TGACCTTTTGGTGGCATCATCAGCTGGGGCCCTGGGCGTGTGGACAGAAGGATGTTCCCGG 855

Qy  1074 TGTGTACACAAAGGTTACCAACTACCTTAGACTGGTCGTGACAACA 1120
           ||||||| ||  ||||||||||||||||||||||||||||||||||
Db   856 TGTGTCACCCAAGTTTACAACTACCTTAGACTACCTAAACCCGATTCGGGACAACA 902
```

RESULT 8
BU845014
LOCUS   BU845014   930 bp   mRNA   linear   EST 16-OCT-2002
DEFINITION AGENCOURT_10412263 NIH_MGC_109 Homo sapiens cDNA clone
            IMAGE:6578533 5', mRNA sequence.
ACCESSION   BU845014
VERSION   BU845014.1 GI:24029455
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2780 row: g column: 13
            High quality sequence stop: 639.
FEATURES   Location/Qualifiers
     source   1..930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6578533"
                     /tissue_type="teratocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_109"
                     /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT   199 a   279 c   266 g   186 t

RESULT 9
BU157720
LOCUS   BU157720   888 bp   mRNA   linear   EST 04-SEP-2002
DEFINITION AGENCOURT_8043192 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084476
            5', mRNA sequence.
ACCESSION   BU157720
VERSION   BU157720.1 GI:22671252
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens

Qy 658 GACATTGCCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCGTGCGTGCCGGAGAGCAGCGTG 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GACATTGCCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGCCGTGCGTGCCGGAGAGCAGCGTG 540

Qy 718 GTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGCAGCGGAGTGTGAG 777
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTCCGCACTGTGTGCCTTCCCCGGCGACCTGCAGCTGCCGGACTGCAGCGGAGTGTGAG 600

Qy 778 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTTCTCCTTTCTATTCGGAGCGGGCTGAAGGAG 837
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CTCTCCGGCTACGGCAAGCATGAGGCCTTGTCCTTCTCCTTTCTATTCGGAGCGGGCTGAAGGAG 660

Qy 838 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACA 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GCTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTACTTAACAGAACA 720

Qy 898 GTCACCGACAACATGCTGT-GTGCTGAGACACTCGG--AGCGGCGGGCCCCAGGCAAAC 954
          |||||||||||||||||| ||||||||||||||| |||||||||||||||||||
Db 721 GTCACCGACAACATGCTGTGGTGCTGGAGACACTCGGGGAGCGGCGGGCCCCAGGCAAAC 780

Qy 955 TTGCACGACGGCCT-GCCAGGGCGATTCGGAGGCGCCCCTGGT-GTGTCTGAACGATGG-C 1011
          ||||||||||||| |||||||||||||||||||||||||||| |||||||||||||||
Db 781 TTGCACGACGCCTGGCCGGGGCGATTCGGAGGCGCCCCTGGTGGGGTCTGAACGATGGCC 840

Qy 1012 CGCATGACTTT--GGTGGGCATCATCAGCTGGGGCCTCGGCGTGG 1055
          |||||||||| |||||||||||||||||||||||||||||||
Db 841 CGCATGACTTTGGGGGGCATCATCAACTGGGGGCCTGGGGGCTGG 886

RESULT 10
BQ927899
LOCUS   BQ927899 AGENCOURT_8775988 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6373489
                         950 bp   mRNA   linear   EST 20-AUG-2002
DEFINITION   AGENCOURT_8775988 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6373489
             5', mRNA sequence.
ACCESSION   BQ927899
VERSION   BQ927899.1   GI:22342930
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 950)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM2550 row: p column: 02
             High quality sequence start: 35
             High quality sequence stop: 684.
FEATURES   Location/Qualifiers
     source   1..950
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6373489"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a

---

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  AUTHORS   1 (bases 1 to 888)
  TITLE   NIH-MGC http://mgc.nci.nih.gov/.
  JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
             Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM2316 row: e column: 21
             High quality sequence stop: 630.
FEATURES   Location/Qualifiers
     source   1..888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6084476"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
                     Site 2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."

BASE COUNT   181 a   265 c   270 g   172 t
ORIGIN

Query Match   69.4%; Score 782.8; DB 13; Length 888;
Best Local Similarity   97.2%; Pred. No. 4.9e-193;
Matches 861; Conservative 0; Mismatches 17; Indels 8; Gaps 6;

Qy 178 GGCAAGGTTTACACAGACACAGAACCCCAGTGCCCAGGCACTGGCCTGGCAAACATAAT 237
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGCAAGGTTTACACAGACACAGAACCCCAGTGCCCAGGCACTGGCCTGGCAAACATAAT 60

Qy 238 TACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 297
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TACTGCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG 120

Qy 298 CTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGTGAGACAGTACAGC 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCACCTGCGGCGTGAGACAGTACAGC 180

Qy 358 CAGCCTCAGTTTGCCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGCCAG 417
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CAGCCTCAGTTTGCCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGCCAG 240

Qy 418 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATA 477
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATA 300

Qy 478 CTCATCAGCTCCTGCTGATTCTCTGCCGCGCCACTGCTTCCAGGAGGAGGTTTCCGCCC 537
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CTCATCAGCTCCTGCTGATTCTCTGCCGCGCCACTGCTTCCAGGAGGAGGTTTCCGCCC 360

Qy 538 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAG 597
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGGCGGAGGAGCAG 420

Qy 598 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 657
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 AAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAT 480

SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 916)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11476 row: k column: 09
            High quality sequence stop: 837.
FEATURES   Location/Qualifiers
    source   1..916
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:5190656"
                 /lab_host="DH10B"
                 /clone_lib="NIH_MGC 116"
                 /note="Organ: pooled colon, kidney, stomach; Vector:
                 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                 source anonymous pool of 3 colons, age 26 yo male, 49 yo
                 female, 71 yo male colon; 46 yo male kidney, and pool of 2
                 stomachs, 62 yo male and 70 yo female. Library is
                 oligo-dT primed and directionally cloned (EcoRV site is
                 destroyed upon cloning). Average insert size 1.4 kb,
                 insert size range 1-3 kb. Library is normalized and
                 enriched for full-length clones and was constructed by C.
                 Gruber (Invitrogen). Research Genetics tracking code
                 023. Note: this is a NIH_MGC Library."
BASE COUNT   204 a   272 c   256 g   184 t
ORIGIN

Query Match   67.8%;   Score 764.4; DB 12; Length 916;
Best Local Similarity 99.7%;   Pred. No. 3.1e-188;
Matches 776;   Conservative   0; Mismatches 1; Indels 1; Gaps 1;

```
Qy    352 TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCC 411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGCCTCCACCCC 60

Qy    412 TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGG 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGG 120

Qy    472 GGCATACTCATCAGCTCCTGCTGATTCTCTCGCTGCTGCTCCTCGCTGCTTCCAGGAGGGTTT 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGCATACTCATCAGCTCCTGCTGATTCTCTCGCTGCTGCTCCTCGCTGCTTCCAGGAGGGTTT 180

Qy    532 CCGCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAG 591
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCTGGCGAGGAG 240

Qy    592 GAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTAC 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGAACACTTAC 300

Qy    652 GACAATGACATTGCCGCTGCTCGAGCTGAAATCGGATTCTCCGGTCTCCCGTGTGCCCAGGAGAGC 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GACAATGACATTGCCGCTGCTCGAGCTGAAATCGGATTCTCCGGTCTCCCGTGTGCCCAGGAGAGC 360

Qy    712 AGCGTGGTTCCGCACTGTGTGCCTTCCCCGGCCGGACTCTGCAGCTGCCCGGAGCTGGACGGAG 771
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGCGTGGTTCCGCACTGTGTGCCTTCCCCGGCCGGACTCTGCAGCTGCCCGGAGCTGGACGGAG 420
```

BASE COUNT   207 a   283 c   279 g   181 t
ORIGIN   NIH_MGC Library."

Query Match   67.8%;   Score 765.2; DB 13; Length 950;
Best Local Similarity 98.1%;   Pred. No. 2e-188;
Matches 806;   Conservative   0; Mismatches 13; Indels 3; Gaps 3;

```
Qy    276 CCACGTGCTGAAGAACCCGAGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCAC 335
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43 CCCCGTGCTGAAGAACCCGAGGCTGACGTGGGAGTACTGTGATGTGCCCTCCTGCTCCAC 102

Qy    336 CTGCGGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGA 395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103 CTGCGGGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGA 162

Qy    396 CATCGCCTCCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGA 455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    163 CATCGCCTCCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGGTCGCCCGGAGA 222

Qy    456 GCGGTTCCTGTGCGGGGGGCCATACTCATCAGCTCCTGCTGGATTCTCTCGCTGCCCCACTG 515
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    223 GCGGTTCCTGTGCGGGGGGCCATACTCATCAGCTCCTGGATTCTCTCGCTGCCCCACTG 282

Qy    516 CTTCCAGGAGAG-TTTCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGG 574
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    283 CTTCCAGGAGAGGTTTCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGG 342

Qy    575 TGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAAT 634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343 TGGTCCCTGGCGGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAAT 402

Qy    635 TCGATGATGACACTTACGACACAATGACATTGCCGCTGCGAGTCGAATCGGATTCGTCCC 694
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 TCGATGATGACACTTACGACACAATGACATTGCCGCTGCGAGTCGAATCGGATTCGTCCC 462

Qy    695 GCTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTCGCCTTCCCCCGGCGGACCTGCAGC 754
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    463 GCTGTGCCCAGGAGAGCAGCGTGGTCCGCACTGTGTCGCCTTCCCCCGGCGGACCTGCAGC 522

Qy    755 TGCCCGGACTGGACGGAGTGTGAGCGTGAGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCTCTT 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    523 TGCCGGACTGGACGGAGTGTGAGCGTGAGCTCTCCGGCTACGGCAAGCATGAGGGCCTTGTCCCTT 582

Qy    815 TCTATTCCGAGCGGCTGAAGGAGCGCTCATGTCAGACTGTACCCATCCAGCCGCCTGCACAT 874
          |||||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db    583 TCTATTCCGAGCGGCTGAAGGAGCGCTCATGTCAGACTGTACCCATCCAGCCGCCTGCACAT 642

Qy    875 CACAACATTTACTTAACAGAACAGTCACCGACACAATGCTGTGTGCTGGAGACACTCGGA 934
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    643 CACAACATTTACTTAACAGAACAGTCACCGACACAATGCTGTGTGCTGGAGACACTCGGA 702

Qy    935 GCGGCGGGGCCCCAGGCAAACTTGCACGACGCGCTGCCAGGGGCGATTCGGGAGGCCCCCTGG 994
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    703 GCGGCGGGGCCCCAGGCAAACTTGCACGACGCGCTGCCAGGGGCGATTCGGGAGGCCCCCTGG 762

Qy    995 TGTGTCTGAACGATGGCCGCATGACCTTTGGTGGCGCATCATCAG--CTGGGGCCTGT 1053
          ||||||||||||||||||||||||||||||||||| ||||||||  |||||||||
Db    763 TGTGTCTGAACGATGGCCGCATGACCTTTGGTGGCGCATCATCAGCCTGGGGCCTGT 822

Qy   1054 GGACAG-AAGGATGTCCCGGGTGTGTACACAAAGGTTACCAA 1094
          |||||| ||| |||||||||||||||||||||||||||||
Db    823 GGACAGAAAGAGAGGCCCGGGTGGGGACCACCAGGGTACCCA 864
```

RESULT 11
BI765247
LOCUS   BI765247   603050338F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190656 5',
            916 bp   mRNA   linear   EST 25-SEP-2001
DEFINITION 603050338F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190656 5'
            mRNA sequence.
ACCESSION   BI765247
VERSION   BI765247.1 GI:15755825
KEYWORDS   EST.

```
Query Match          62.7%;   Score 706.8;  DB 10;   Length 948;
Best Local Similarity 93.3%;   Pred. No. 3.2e-173;
Matches  793;  Conservative  0;  Mismatches  52;  Indels  5;  Gaps  5;

Qy    72 GGAAACAGTGACTGCTACT-TTGGGAATGGTCAGCCTACCGTGGCACGCACAGCCTCA 130
         |||||||||| |||||||||  |||||||||||||||||||||||||||||||||||||
Db     8 GAGGAACAGTGACTGCTATTATTGGGAATGGTCAGCCTACCGTGGCACGCACAGCCTCA 67

Qy   131 CCGAGTCGGTGCCTCCTGCCTCCC-GTGGAATTCATGATCCTGATAGGCAAGGTTTAC 189
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    68 CCGAGTCGGTGCCTCCTGCCTCCCGGTGGAATTCCATGATCCTGATAGGCAAGGTTTAC 127

Qy   190 ACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAAT 249
         |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   128 ACAGCACAGAACCCAGTGCCCAGGCACTGGGCCTGGGCAAACATAATTACTGCCGGAAT 187

Qy   250 CCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAG 309
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 CCTGATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAG 247

Qy   310 TACTGTGATGTGCCCTCCTGCTCCACCTGGCGGCTGAGACAGCAGCAGCCTCAGTTT 369
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TACTGTGATGTGCCCTCCTGCTCCACCTGGCGGCTGAGACAGCAGCAGCCTCAGTTT 307

Qy   370 CGCATCAAAGGAGGAGGCTCTTCGCCGACATCGCCTCCGCCCCCTGGCAGGCTGCCATCTTT 429
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 CGCATCAAAGGAGGAGGCTCTTCGCCGACATCGCCTCCGCCCCCTGGCAGGCTGCCATCTTT 367

Qy   430 GCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGCTCC 489
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 GCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGCATACTCATCAGCTCC 427

Qy   490 TGGTGGATTCTCTCTGCTTCCGCCCCACTGCTTCCGCCCCACCACCTGACG 549
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 TGCTGGATTCTCTCTGCTTCCGCCCCACTGCTTCCGCCCCACCACCTGACG 487

Qy   550 GTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGGCAGAAATTTGAAGTC 609
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 GTGATCTTGGGCAGAACATACCGGTGGTCCCTGGCGAGGAGGCAGAAATTTGAAGTC 547

Qy   610 GAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTG 669
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTGCGCTG 607

Qy   670 CTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGAGCAGCAGCGTGGTTCCGCACTGTG 729
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 CTGCAGCTGAAATCGGATTCGTCCCGCGTGCCCAGGAGAGCAGCAGCGTGGTTCCGCACTGTG 667

Qy   730 TGCCTTCCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGAGCTGTCTCCGGCTAC 789
         |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 TGCCTT-CCCCGGCGGACCTGCAGCTGCCGGACTGGACGGAGTGAGCTGTCTCCGGCTAC 726

Qy   790 GGCAAGCATGAGGCCCTTGTCTCCTTTCATTCGGAGCGGCTGAAGGAGGCTCATGTCAGA 849
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 GGCAAGCATGAGG-CTTGCTCCTTTCATTCGGAGCGGCTGAAGGGGCTCATGTCGGA 785

Qy   850 CTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAC 909
         ||| |||||||||||||||||||||| ||| |||||| |||  |  |||||||||||||
Db   786 CTG-GTCCATCCAGCCGCTGGGCTTCCCCCACCTTTCTTCAGGACGACAGCGGCGCCTGC 844

Qy   910 ATGCTGTGTG 919
         ||| || |||
Db   845 TGGGGCTGGG 854
```

RESULT 13
BG763582

LOCUS   BG763582   958 bp   mRNA   linear   EST 15-MAY-2001
DEFINITION   602735987F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4861268 5',
             mRNA sequence.
ACCESSION   BG763582
VERSION   BG763582.1 GI:14074235
KEYWORDS   EST.

```
Qy   772 TGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCCTTGTCTCCTTTCATTCGGAGCGGCTG 831
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TGTGAGCTCTCCGGCTACGGCAAGCATGAGGCCCTTGTCTCCTTTCATTCGGAGCGGCTG 480

Qy   832 AAGGAGGGCTCATGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAAC 891
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGGAGGGCTCATGTCAGACTGTCAGACTGTACCATCCAGCCGCTGCACATCACAACATTTACTTAAC 540

Qy   892 AGAACAGTCACCGACAACATGCTGTGTGTGGAGACACTCGGAGCGGCGGGGCCCCAGGCA 951
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AGAACAGTCACCGACAACATGCTGTGTGTGGAGACACTCGGAGCGGCGGGGCCCCCAGGCA 600

Qy   952 AACTTGACGACGCCTGCCAGGGGCGATTCGGAGGCGACCCCCTGGTGTGTGAACGATGGC 1011
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AACTTGACGACGCCTGCCAGGGGCGATTCGGAGGCGACCCCCTGGTGTGTGAACGATGGC 660

Qy  1012 CGCATGACTTGGTGGGCATCATCAGCTGGGGGCTGGGGCTGTGGACAGAAGGATGTCCCG 1071
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGCATGACTTGGTGGGCATCATCAGCTGGGGGCTGGGGCTGTGGACAGAAGGATGTCCCG 720

Qy  1072 GGTGTGTACACAAAGGTTACCAACTACCTAGAC-TGGATTCGTGACAACATGCGACCG 1128
         |||||||| ||||||||  |||||||||||||| |||||||||||||||||||||||||
Db   721 GGTGTGTACACCAAGGTTACCAACTACCTAGACTTGGATTCGTGACAACATGCGACCG 778
```

RESULT 12
BE616613

LOCUS   BE616613   948 bp   mRNA   linear   EST 24-AUG-2000
DEFINITION   601278864F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3611007 5',
             mRNA sequence.
ACCESSION   BE616613
VERSION   BE616613.1 GI:9898212
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM267 row: p column: 16
             High quality sequence start: 2
             High quality sequence stop: 759.
FEATURES   Location/Qualifiers
     source   1..948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3611007"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_39"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California, Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life
                     Technologies)."
BASE COUNT   173 a   295 c   293 g   187 t
ORIGIN
```

```
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 958)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTP
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM1719 row: n  column: 21
              High quality sequence stop: 810.
FEATURES      Location/Qualifiers
     source   1..958
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4861268"
              /tissue_type="melanotic melanoma, high MDR (cell line)"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_49"
              /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
              EcoRI; cDNA made by oligo-dT priming. Directionally cloned
              into EcoRI/XhoI sites using the following 5' adaptor:
              GGCACGAG(G). Size-selected >500bp for average insert size
              1.8kb. Library constructed by Ling Hong in the laboratory
              of Gerald M. Rubin (University of California, Berkeley)
              using ZAP-cDNA synthesis kit (Stratagene) and Superscript
              II RT (Life Technologies). Note: this is a NIH_MGC
              Library."
BASE COUNT    198 a   285 c   282 g   193 t
ORIGIN

Query Match               Score 700.4;  DB 12;  Length 958;
Best Local Similarity   62.1%;  Pred. No. 1.5e-171;
Matches   840;  Conservative   0;  Mismatches   41;  Indels   15;  Gaps   10;

Qy    56  CCCAGGCGGCCCTCGAGGGAAACAGTGACTGCTACTTTGGGATGGTCAGCTACCGTG 115
          || ||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CCCCTGCCTCTGAGGGAAACAGTGACTGCTACTTTGGGATGGTCAGCTACCGTG 120

Qy   116  GCACGCACAGCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GCACGCACAGCCCTCACCGAGTCGGGTGCCTCCTGCCTCCCGTGGAATTCCATGATCCTGA 180

Qy   176  TAGGCAAGGTTTACACAGCACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TAGGCAAGGTTTACACAGCACAGCACAGAACCCCAGTGCCCAGGCACTGGGCCTGGGCAAACATA 240

Qy   236  ATTACTGCCGGAATCCTGATGGGGATGGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  ATTACTGCCGGAATCCTGATGGGGATGGGCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 300

Qy   296  GGCTGACGTGGGAGTACTGTGATGTGCCCCTGCTCCACCTGCGCGGCCTGAGACAGTACA 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GGCTGACGTGGGAGTACTGTGATGTGCCCCTGCTCCACCTGCGCGGCCTGAGACAGTACA 360

Qy   356  GCCAGCCTCAGTTTGCCATCAAAGGAGGGCTCTTCGCCGACATCGCCTTCCCACCCCTGGC 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GCCAGCCCTCAGTTTGCCATCAAAGGAGGGCTCTTCGCCGACATCGCCTTCCCCACCCCTGGC 420

Qy   416  AGGCTGCCATCTTTGCCAAGCACCAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AGGCTGCCATCTTTGCCAAGCACCAGGAGGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCA 480

Qy   476  TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCCACTGCTTCTTCCAGGAGAGGTTCCGC 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TACTCATCAGCTCCTGCTGGATTCTCTCGCCGCCCCACTGCTTCTTCCAGGAGAGGTTCCTG 540

Db   536  CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGAGGAGGAGC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541  CCCACCACTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCCTGCGGAGGAGAGC 600

Qy   596  AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 660

Qy   656  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGT--CCGGTCGTGCCCAGGAGCAGC 714
          |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db   661  ATGACATTGCGCTGCTGCAGCTGAAATCGGATTCGTCCGTCCCCGGTCGTGCCCAGGAGCAGC 720

Qy   715  -GTGGTCCGGCACTGTGTGCCTT--CCCCGGCGGCGACCTGCAGCTGCCGGGGACTGGAC-GGA 770
           |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||| |||
Db   721  GGTGGTCCGGCACTGTGTGCCTTTCCCCGGCGGCGACCTGCAGCTGCCGGGGACTGGACTGGGGA 780

Qy   771  GTGTGAGCTGTCCG-GCTACGGCAAGCAT--GAGGC 827
          |||||||||||||||  ||||||||||||  ||||||
Db   781  GTGTGAGCTGTCCGAGCTACGGCAAGCATTGAGGGCCTGGCCTTGTCTCCTTTCTATTCGGAGCG 840

Qy   828  GCTGAAGAGGCTTCATGTCAGACTGTACCCATCCAGCCGCTGCACATCACAACATTTACT 887
          |||||||||||||||||||||||| |||||  ||||||||||||||||||||||||||||||
Db   841  G-TGAAGGAGGGCTCATG--CGACTGGACCCCATTCCACCGGTGGCCATTCACAACTTTACT 897

Qy   888  TAACAGAACAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGGC 943
          |   |||||||||||||||||   |||||||||||||||||||||||||||||||||
Db   898  T-ACAGAACAGTACC---GAAAATGGTGTGTGCGGGACCCCGACGGGGGGCCAGGC 949
```

RESULT 14
```
BQ960216
LOCUS       BQ960216   AGENCOURT_8945604 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6459629
            5', mRNA sequence.                      974 bp    mRNA    linear    EST 21-AUG-2002
DEFINITION  AGENCOURT_8945604 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6459629
            5', mRNA sequence.
ACCESSION   BQ960216
VERSION     BQ960216.1  GI:22375694
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 974)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2645  row: a   column: 06
            High quality sequence stop: 603.
FEATURES    Location/Qualifiers
     source 1..974
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6459629"
            /tissue_type="epidermoid carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_101"
            /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
            XhoI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 709)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-r@mail.nih.gov
    cDNA Library Preparation: Ling Hong/Rubin Laboratory
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
    Plate: LLCM752 row: d column: 17
    High quality sequence start: 3
    High quality sequence stop: 701.

FEATURES   Location/Qualifiers
source   1..709
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:3925600"
    /tissue_type="choriocarcinoma"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH_MGC_21"
    /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
    Site_2: EcoRI; cDNA made by oligo-dT priming.
    Directionally cloned into EcoRI/XhoI sites using the
    following 5' adaptor: GGCACGAG(G). Size-selected >500bp
    for average insert size 1.8kb. Library constructed by
    Ling Hong in the laboratory of Gerald M. Rubin (University
    of California, Berkeley) using ZAP-cDNA synthesis kit
    (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT   147 a   219 c   203 g   140 t
ORIGIN

Query Match   61.5%;   Score 693.4;   DB 10;   Length 709;
Best Local Similarity   99.7%;   Pred. No. 8.9e-170;
Matches   705;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

Qy   263 CCAAGCCCTGGTTCGCCACGTGCTGAAGAACCCGCAGGCTGCCCAGCGTCGGGAGTACTGTGATGTGC 322
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCAAGCCCTGGTTCGCCACGTGCTGAAGAACCCGCAGGCTGCCCAGCGTCGGGAGTACTGTGATGTGC 60

Qy   323 CCTTCGTTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTCGCATCAAAGGAG 382
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCTCCGTTCCACCTGCGGCCTGAGACAGTACAGCCAGCCTCAGTTCGCATCAAAGGAG 120

Qy   383 GGCTCTTCGCGACATGCGCTGCTCCCACCCCTGCGAGGCTGCCCATCTTTGCCAAGCACCAGA 442
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGCTCTTCGCGACATGCGCTGCTCCCACCCCTGCGAGGCTGCCCATCTTTGCCAAGCACCAGA 180

Qy   443 GGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTACCTGGGGCATACTACCTCGTCGGATTCTCT 502
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGTCGCCCGGAGAGCGGTTCCTGTGCGGGGGCATACTACCTGGGGCATACTACCTCGTCGGATTCTCT 240

Qy   503 CTGCCGCCACTGCTTCCAGGAGAGGTTCCGCCCCCACCACCTGACGGTGATCTTGGGCA 562
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGCCGCCACTGCTTCCAGGAGAGGTTCCGCCCCCACCACCTGACGGTGATCTTGGGCA 300

Qy   563 GAACATACCGGGTGGTCCCTGGCCCTGGCCAGGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTG 622
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAACATACCGGGTGGTCCCTGGCCCTGGCCAGGAGGAGGAGCAGAAATTTGAAGTCGAAAATACATTG 360

Qy   623 TCCATAAGGAATTCGATGATGACAGTTACGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAT 682
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCATAAGGAATTCGATGATGACAGTTACGACACTTACGACAATGACATTGCGCTGCTGCAGCTGAAT 420

Qy   683 CGGATTCGTCCCGCTGTGCCCGTGGTCCGGAGAGCAGCCGTGGTCCACTGTGTGCCTTCCCCGG 742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGGATTCGTCCCGCTGTGCCCGTGGTCCGGAGAGCAGCCGTGGTCCACTGTGTGCCTTCCCCGG 480

(left column)

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT   215 a   293 c   279 g   187 t
ORIGIN

Query Match   61.6%;   Score 694.4;   DB 13;   Length 974;
Best Local Similarity   97.8%;   Pred. No. 5.5e-170;
Matches   757;   Conservative   0;   Mismatches   11;   Indels   6;   Gaps   5;

Qy   342 CCTTGAGACAGTACAGCCAGCCTCAGTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGC 401
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   19 CCTTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTCGCCGACATCGC 78

Qy   402 CTCCCACCCCTGCGAGGCTGCCATCTTTGCCAAGCACCAGGAGGAGCGCCCGGAGAGCCGGTT 461
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 CTCCCACCCCTGCGAGGCTGCCATCTTTGCCAAGCACCAGGAGGAGCGCCCGGAGAGCCGGTT 138

Qy   462 CCTGTGCGGGGGCATACTCATCAGCTCCTGGATTCTCTCTCGCCGCCCACTGCTTCCA 521
           |||||||||||||||||||| |||||||||||||||||||| |||||||||||||||||
Db   139 CCTGTGCGGGGGCATACTACTCATCAGCTCCTGGATTCTCTCTCGCCGCTC-CTGCTTCCA 197

Qy   522 GGAGAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCC 581
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 GGAGAGGTTTCCGCCCCACCACCTGACGGTGATCTTGGGCAGAACATACCGGGTGGTCCC 257

Qy   582 TGGCCGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTTCCATAAGGAATTCGATGA 641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 TGGCCGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTGTTCCATAAGGAATTCGATGA 317

Qy   642 TGACACTTACGACAATGACATTGCGCTGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGC 701
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 TGACACTTACGACAATGACATTGCGCTGCTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGC 377

Qy   702 CCAGGAGAGCAGCCGTGGTCCGGACACTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGA 761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 CCAGGAGAGCAGCCGTGGTCCGGACACTGTGCCTTCCCCGGCGGACCTGCAGCTGCCGGA 437

Qy   762 CTGGGACCGGAGTGTGAGCTCCGGCAAGCATGCGGACCTGTGCCTTGTTGTCCTTTCTATTC 821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTGGGACCGGAGTGTGAGCTCCGGCAAGCATGCGGACCTGTGCCTTGTTGTCCTTTCTATTC 497

Qy   822 GGAGCGGCTGAAGGAGGCTCATGTGACTGTACCCATCCAGCCGCTGCACATCACAACA 881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498 GGAGCGGCTGAAGGAGGCTCATGTGACTGTACCCATCCAGCCGCTGCACATCACAACA 557

Qy   882 TTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGG 941
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 TTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTGAGACACTCGGAGCGGCGGG 617

Qy   942 GCCCCAGGCAAACTTGACGACGACCGCTGCCAGGCGAGTTCGGAGG--CCCCTGGTGTGTC 1000
           ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   618 GCCCCAGGCAAACTTGACGACGACCGCTGCCAGGCGAGTTCGGAGGGCCCCCTGGTGTGTC 677

Qy   1001 TGAACGATGGCCCGCAT-GACTTTGGTGGGCATCATCAGCT-GGGGCCTGGGCGTGTGGACA 1058
           ||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||
Db   678 TGAACGATGGCCCGCATGGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCGTGTGGACA 737

Qy   1059 GAAGGATGTCCC--GGGTGTGTACACAAAGGTTACCAACTACCTAGACTGGATT 1110
           ||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   738 GAAGGATGTCCCGGGGTGTGTACACAAAGGTTACCAACTACCTAAACTGGAAT 791

RESULT 15
BE732704
LOCUS   BE732704   709 bp   mRNA   linear   EST 15-SEP-2000
DEFINITION   601571366F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925600 5',
    mRNA sequence.
ACCESSION   BE732704
VERSION   BE732704.1   GI:10146696
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens

```
Qy    743 CGGACCTGCAGCTGCCGGACTGACGGAGTGTGAGCTTCCCGGCTACGGCAAGCATGAGG  802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGGACCTGCAGCTGCCGGACTGGACGGAGTGTGAGCTTCTCCGGCTACGGCAAGCATGAGG  540

Qy    803 CCTTGTCTCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCA  862
          ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTTGTCCCCTTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTCAGACTGTACCCATCCA  600

Qy    863 GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG  922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCCGCTGCACATCACAACATTTACTTAACAGAACAGTCACCGACAACATGCTGTGTGCTG  660

Qy    923 GAGACACTCGGAGCGGCGGCGGGCCCCAGGCAAACTTGCACGACGCCTGC  969
          |||||||||||||||||||||||| |||||||||||||||||||||||||
Db    661 GAGACACTCGGAGCGGCGGCGGG-CCCAGGCAAACTTGCACGACGCCTGC  706
```

Search completed: August 19, 2003, 08:29:08
Job time : 3305.95 secs